GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:37:13 ; Search time 52.2391 Seconds
(without alignments)
48.679 Million cell updates/sec

Title: US-10-048-209-2
Perfect score: 52
Sequence: 1 KQYTSIHHG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	11	4	AAB67792	Aab67792 Cytoplasm
2	52	100.0	12	4	AAB67793	Aab67793 Cytoplasm
3	52	100.0	13	4	AAB67794	Aab67794 Cytoplasm
4	52	100.0	15	4	AAB67798	Aab67798 Cytoplasm
5	52	100.0	19	7	ADC49255	Adc49255 Human inh
6	52	100.0	26	2	AAR58933	Aar58933 KQYTSI-ad
7	52	100.0	32	6	AAO19883	Aao19883 Human amy
8	52	100.0	33	7	ADC49254	Adc49254 Human inh
9	52	100.0	41	4	AAM16658	Aam16658 Peptide #
10	52	100.0	41	4	ABB35642	Abb35642 Peptide #
11	52	100.0	41	4	AAM29142	Aam29142 Peptide #
12	52	100.0	41	4	ABB30475	Abb30475 Peptide #
13	52	100.0	41	4	ABB21071	Abb21071 Protein #
14	52	100.0	41	4	AAM56458	Aam56458 Human bra
15	52	100.0	41	4	AAM04374	Aam04374 Peptide #
16	52	100.0	41	5	ABG38416	Abg38416 Human pep
17	52	100.0	44	2	AAW53985	Aaw53985 Human ALZ
18	52	100.0	47	2	AAR58917	Aar58917 Cytoplasm
19	52	100.0	47	2	AAW26395	Aaw26395 Amyloid p
20	52	100.0	47	2	AAW26402	Aaw26402 Amyloid p
21	52	100.0	47	2	AAW26400	Aaw26400 Amyloid p
22	52	100.0	47	2	AAW26403	Aaw26403 Amyloid p
23	52	100.0	47	2	AAW26399	Aaw26399 Amyloid p
24	52	100.0	47	2	AAW26401	Aaw26401 Amyloid p
25	52	100.0	47	2	AAW26520	Aaw26520 Amyloid p
26	52	100.0	47	2	AAW26518	Aaw26518 Amyloid p
27	52	100.0	47	2	AAW26521	Aaw26521 Amyloid p
28	52	100.0	47	2	AAW26519	Aaw26519 Amyloid p
29	52	100.0	47	2	AAW26513	Aaw26513 Amyloid p
30	52	100.0	47	2	AAW26517	Aaw26517 Amyloid p
31	52	100.0	47	2	AAW42984	Aaw42984 APP isofo
32	52	100.0	47	2	AAW42986	Aaw42986 APP isofo
33	52	100.0	47	2	AAW42987	Aaw42987 APP isofo
34	52	100.0	47	2	AAW42983	Aaw42983 APP isofo
35	52	100.0	47	2	AAW42985	Aaw42985 APP isofo
36	52	100.0	47	2	AAW44755	Aaw44755 APP-REP 7
37	52	100.0	47	2	AAW44753	Aaw44753 APP-REP 7
38	52	100.0	47	2	AAW44749	Aaw44749 APP-REP 7
39	52	100.0	47	2	AAW44756	Aaw44756 APP-REP 7
40	52	100.0	47	2	AAW44754	Aaw44754 APP-REP 7
41	52	100.0	47	2	AAW44757	Aaw44757 APP-REP 7
42	52	100.0	47	2	AAW44274	Aaw44274 Amyloid p
43	52	100.0	47	4	AAB67791	Aab67791 Cytoplasm
44	52	100.0	47	7	ADC49250	Adc49250 Human inh
45	52	100.0	47	7	ADC49249	Adc49249 Inhibitor

ALIGNMENTS

RESULT 1
AAB67792
ID AAB67792 standard; peptide; 11 AA.
XX
AC AAB67792;
XX
DT 11-JUN-2001 (first entry)
XX
DE Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1

```
FT                   /note= "this residue represents an internalisation
FT                   peptide such as the sequence given in AAB67795"
FT   Misc-difference 11
FT                   /note= "this residue represents V, VV, VVE, VVEV, VVEVD"
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Claim 1; Page 13; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 11 AA;

  Query Match             100.0%;  Score 52;  DB 4;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.0033;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db          2 KQYTSIHHG 10


RESULT 2
AAB67793
ID   AAB67793 standard; peptide; 12 AA.
XX
AC   AAB67793;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW   Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
```

```
FT   Misc-difference 1
FT                   /note= "this residue represents an internalisation
FT                   peptide such as the sequence given in AAB67795"
FT   Misc-difference 12
FT                   /note= "this residue represents V, VV, VVE, VVEV, VVEVD"
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Claim 1; Page 13; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 12 AA;

  Query Match             100.0%;  Score 52;  DB 4;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.0036;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db          3 KQYTSIHHG 11


RESULT 3
AAB67794
ID   AAB67794 standard; peptide; 13 AA.
XX
AC   AAB67794;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW   Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "this residue represents an internalisation
FT                   peptide such as the sequence given in AAB67795"
FT   Misc-difference 13
FT                   /note= "this residue represents V, VV, VVE, VVEV, VVEVD"
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Claim 1; Page 13; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 52;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.0039;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db          4 KQYTSIHHG 12


RESULT 4
AAB67798
ID   AAB67798 standard; peptide; 15 AA.
XX
AC   AAB67798;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW   Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Disclosure; Page 2; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 52;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.0046;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db          3 KQYTSIHHG 11


RESULT 5
ADC49255
ID   ADC49255 standard; protein; 19 AA.
XX
AC   ADC49255;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human inhibitor of metabolic degradation of APP variant #4.
XX
KW   human; inhibitory factor; metabolic degradation of APP;
KW   amyloid precursor protein; Alzheimer's disease; mutant; mutein.
XX
OS   Homo sapiens.
XX
PN   JP2002360252-A.
XX
PD   17-DEC-2002.
XX
```

PF 27-APR-2001; 2001JP-00133178.
XX
PR 27-APR-2001; 2001JP-00133178.
XX
PA (SUZU/) SUZUKI T.
PA (SUMJ) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2003-516151/49.
XX
PT An inhibitory factor of metabolic degradation of amyloid precursor
PT protein (APP) which inhibits formation of beta-amyloid, useful in the
PT treatment of Alzheimer's disease.
XX
PS Disclosure; SEQ ID NO 15; 33pp; Japanese.
XX
CC The invention relates to an inhibitory factor of metabolic degradation of
CC APP. The factor is useful in the treatment of Alzheimer's disease. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 19 AA;

  Query Match   100.0%; Score 52; DB 7; Length 19;
  Best Local Similarity 100.0%; Pred. No. 0.006;
  Matches   9; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQYTSIHHG 9
              |||||||||
Db   3 KQYTSIHHG 11

RESULT 6
AAR58933
ID AAR58933 standard; peptide; 26 AA.
XX
AC AAR58933;
XX
DT 25-MAR-2003 (revised)
DT 15-APR-1995 (first entry)
XX
DE KQYTSI-added variant of His657-Lys676 region of human cytoplasmic amyloid
DE precursor protein APP695.
XX
KW Amyloid precursor protein; isoform APP 695; beta amyloid; peptide 20; Go;
KW GTP binding protein; couplone region; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO9419692-A1.
XX
PD 01-SEP-1994.
XX
PF 17-FEB-1994; 94WO-US001712.
XX
PR 18-FEB-1993; 93US-00019208.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Nishimoto I;
XX
DR WPI; 1994-294486/36.
XX
PT Identifying cpds. useful for treating or preventing Alzheimer's disease -
PT by determining whether it interferes with the association of the couplone
PT portion of amyloid precursor protein to G polypeptide.
XX
PS Disclosure; Page 34; 71pp; English.
XX
CC Beta amyloid is synthesised as part of a larger protein referred to as
CC amyloid precursor protein (APP), which has a number of isoforms in
CC humans, including APP695 and APP770. The amino terminal of beta amyloid
CC is generated by cleavage of a peptide bond of APP which in APP695 lies
CC between Met596 and Asp597. APP forms a complex with Go, a GTP-binding
CC protein (or "G protein") in brain. The cytoplasmic App695 sequence His657
CC -Lys676 (AAR58913) possesses a specific Go-activating function, and is
CC necessary for complex formation of this APP with Go. This sequence
CC (AAR58913), sometimes referred to as the "couplone" region of APP, is
CC completely conserved in APP751 and App770, as well as in mouse APP695,
CC and suggests that abnormal APP-Go signalling is involved in the
CC Alzheimer's disease process. This peptide is referred to a 'peptide 20'.
CC Variants of peptide 20 are given in AAR58930-R58934. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 26 AA;

  Query Match   100.0%; Score 52; DB 2; Length 26;
  Best Local Similarity 100.0%; Pred. No. 0.0084;
  Matches   9; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQYTSIHHG 9
              |||||||||
Db   1 KQYTSIHHG 9

RESULT 7
AAO19883
ID AAO19883 standard; peptide; 32 AA.
XX
AC AAO19883;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human amyloid precursor protein APP immunogenic peptide #3.
XX
KW Human; APP; amyloid precursor protein; immunogen; Alzheimer's disease;
KW high-throughput screening; neuroprotective; nootropic; antiparkinsonian.
XX
OS Homo sapiens.
XX
PN WO2003001881-A2.
XX
PD 09-JAN-2003.
XX
PF 26-JUN-2002; 2002WO-US020267.
XX

PR 26-JUN-2001; 2001US-0300959P.
XX
PA (NYME-) NEW YORK STATE OFFICE MENTAL HEALTH.
XX
PI Mathews PM, Nixon RA, Schmidt SD, Jiang Y;
XX
DR WPI; 2003-210182/20.
XX
PT Identifying compounds that modulates the generation of metabolites
PT associated with a disease or disorder, for treating e.g. Alzheimer's
PT disease by determining levels of a cellular component protein, or its
PT conformation state.
XX
PS Example 1; Page 29; 69pp; English.
XX
CC The present invention relates to a method of identifying compounds that
CC modulate the generation of one or more metabolites associated with a
CC disease or disorder comprising determining levels of a cellular component
CC protein or a conformation state of a cellular precursor protein. In
CC particular, the method can be used to determine levels of amyloid
CC precursor protein (APP), which is associated with Alzheimer's disease. It
CC is also useful for identifying compounds as drugs for treating diseases
CC or disorders associated with metabolic and/or proteolytic pathways, e.g.
CC Alzheimer's disease, Parkinson's disease, Huntington's disease, lysosomal
CC storage disorders, prion diseases, the tau-based neurodegenerative
CC disorders, and other non-AD amylcidoses. The present sequence is an
CC immunogenic portion of human APP
XX
SQ Sequence 32 AA;

  Query Match   100.0%; Score 52; DB 6; Length 32;
  Best Local Similarity   100.0%; Pred. No. 0.011;
  Matches   9; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQYTSIHHG 9
              |||||||||
Db   7 KQYTSIHHG 15

RESULT 8
ADC49254
ID ADC49254 standard; protein; 33 AA.
XX
AC ADC49254;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human inhibitor of metabolic degradation of APP variant #3.
XX
KW human; inhibitory factor; metabolic degradation of APP;
KW amyloid precursor protein; Alzheimer's disease; mutant; mutein.
XX
OS Homo sapiens.
XX
PN JP2002360252-A.
XX
PD 17-DEC-2002.
XX
PF 27-APR-2001; 2001JP-00133178.
XX
PR 27-APR-2001; 2001JP-00133178.
XX
PA (SUZU/) SUZUKI T.
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2003-516151/49.
XX
PT An inhibitory factor of metabolic degradation of amyloid precursor
PT protein (APP) which inhibits formation of beta-amyloid, useful in the
PT treatment of Alzheimer's disease.
XX
PS Disclosure; SEQ ID NO 14; 33pp; Japanese.
XX
CC The invention relates to an inhibitory factor of metabolic degradation of
CC APP. The factor is useful in the treatment of Alzheimer's disease. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 33 AA;

  Query Match   100.0%; Score 52; DB 7; Length 33;
  Best Local Similarity   100.0%; Pred. No. 0.011;
  Matches   9; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQYTSIHHG 9
              |||||||||
Db   3 KQYTSIHHG 11

RESULT 9
AAM16658
ID AAM16658 standard; protein; 41 AA.
XX
AC AAM16658;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #3092 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 27; SEQ ID NO 21484; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 52; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQYTSIHHG 9
|||||||||
Db 30 KQYTSIHHG 38

RESULT 10
ABB35642
ID ABB35642 standard; peptide; 41 AA.
XX
AC ABB35642;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #3148 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 28277; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 52; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQYTSIHHG 9
|||||||||
Db 30 KQYTSIHHG 38

RESULT 11
AAM29142
ID AAM29142 standard; protein; 41 AA.
XX
AC AAM29142;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #3179 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX

PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 29411; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 52; DB 4; Length 41;
  Best Local Similarity   100.0%; Pred. No. 0.014;
  Matches   9; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQYTSIHHG 9
              |||||||||
Db   30 KQYTSIHHG 38

RESULT 12
ABB30475
ID ABB30475 standard; peptide; 41 AA.
XX
AC ABB30475;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #3126 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 13443; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 52; DB 4; Length 41;
  Best Local Similarity   100.0%; Pred. No. 0.014;
  Matches   9; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQYTSIHHG 9
              |||||||||
Db   30 KQYTSIHHG 38

RESULT 13
ABB21071
ID ABB21071 standard; protein; 41 AA.
XX

AC ABB21071;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #3070 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 22841; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 52; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQYTSIHHG 9
|||||||||
Db 30 KQYTSIHHG 38

RESULT 14
AAM56458
ID AAM56458 standard; protein; 41 AA.
XX
AC AAM56458;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28563.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 28563; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 52; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.014;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQYTSIHHG 9
              |||||||||
Db         30 KQYTSIHHG 38


RESULT 15
AAM04374
ID   AAM04374 standard; protein; 41 AA.
XX
AC   AAM04374;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Peptide #3056 encoded by probe for measuring breast gene expression.
XX
KW   Probe; human; breast disease; breast cancer; development disorder;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US000661.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression in
PT   a human breast.
XX
PS   Claim 27; SEQ ID NO 13114; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes
CC   (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC   such probe. The probes are useful for measuring human gene expression in
CC   a human breast sample, where the probe hybridises at high stringency to a
CC   nucleic acid expressed in the human breast. The probes are useful for
CC   predicting, diagnosing, grading, staging, monitoring and prognosing
CC   diseases of the human breast, particularly those diseases with polygenic
CC   aetiology. The diseases include: breast cancer, disorders of development,
CC   inflammatory diseases of the breast, fibrocystic changes, proliferative
CC   breast disease and non-carcinoma tumours. Note: The sequence data for
CC   this patent did not form part of the printed specification, but was
CC   obtained in electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 41 AA;

  Query Match             100.0%;  Score 52;  DB 4;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 0.014;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db         30 KQYTSIHHG 38


Search completed: October  4, 2004, 18:52:36
Job time : 55.2391 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:45:55 ; Search time 13.3043 Seconds
(without alignments)
65.071 Million cell updates/sec

Title: US-10-048-209-2
Perfect score: 52
Sequence: 1 KQYTSIHHG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	82	2	PQ0438	Alzheimer's diseas
2	52	100.0	695	1	A49795	Alzheimer's diseas
3	52	100.0	695	2	A27485	Alzheimer's diseas
4	52	100.0	695	2	S00550	Alzheimer's diseas
5	52	100.0	770	1	QRHUA4	Alzheimer's diseas
6	49	94.2	747	2	JH0773	Alzheimer's diseas
7	39	75.0	452	2	S58994	NADH2 dehydrogenas
8	38	73.1	488	2	S64140	hypothetical prote
9	37	71.2	442	2	AH1661	ATP-dependent RNA
10	37	71.2	2584	2	T24158	hypothetical prote
11	37	71.2	2606	2	T24157	hypothetical prote
12	36	69.2	275	2	C75529	competence protein
13	36	69.2	412	2	T24023	hypothetical prote
14	36	69.2	443	2	T46532	probable dTDP-4-ke
15	35	67.3	150	2	F96924	flavodoxin [import
16	35	67.3	264	2	T15289	hypothetical prote
17	35	67.3	443	2	B39794	transcription fact
18	35	67.3	444	1	A39794	transcription fact
19	35	67.3	739	2	AG1667	phosphoribosylform
20	35	67.3	739	2	AI1295	phosphoribosylform
21	35	67.3	1268	2	B88209	protein K02A2.6 [i
22	35	67.3	1313	2	T29193	hypothetical prote
23	34	65.4	99	2	B82524	hypothetical prote
24	34	65.4	254	2	T29556	hypothetical prote
25	34	65.4	332	2	T25779	hypothetical prote
26	34	65.4	393	2	C96592	hypothetical prote
27	34	65.4	410	2	AF1660	aminopeptidases ho
28	34	65.4	410	2	AG1288	aminopeptidases ho
29	34	65.4	637	2	H70535	hypothetical prote
30	34	65.4	784	2	T01139	Mutator-like trans
31	34	65.4	848	2	E85087	hypothetical prote
32	33	63.5	117	2	JT0560	hemerythrin beta c
33	33	63.5	117	2	S50178	hemerythrin beta c
34	33	63.5	142	2	T10871	y4kQ protein - Rhi
35	33	63.5	174	2	A47113	glucuronosyltransf
36	33	63.5	216	2	E97240	transcription regu
37	33	63.5	339	2	B34895	transcription fact
38	33	63.5	344	1	TWXL3	transcription fact
39	33	63.5	361	2	D83798	phosphoserine amin
40	33	63.5	406	2	S76307	hypothetical prote
41	33	63.5	434	2	C95362	probable aminopept
42	33	63.5	530	2	C47113	glucuronosyltransf
43	33	63.5	530	2	S68200	glucuronosyltransf
44	33	63.5	588	2	T04736	hypothetical prote
45	33	63.5	645	2	A83054	acetyl-coenzyme A

ALIGNMENTS

RESULT 1
PQ0438
Alzheimer's disease amyloid A4 protein precursor - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C;Accession: PQ0438; C60045
R;Davidson, J.S.; West, R.L.; Kotikalapudi, P.; Maroun, L.E.
Biochem. Biophys. Res. Commun. 188, 905-911, 1992
A;Title: Sequence and methylation in the beta/A4 region of the rabbit amyloid precursor protein gene.
A;Reference number: PQ0438; MUID:93075180; PMID:1445331
A;Accession: PQ0438
A;Molecule type: DNA
A;Residues: 1-82 <DAV>
A;Cross-references: GB:M83558; GB:M83657
R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157
A;Accession: C60045
A;Molecule type: mRNA
A;Residues: 12-68 <JOH>
A;Cross-references: EMBL:X56129
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; Alzheimer's disease; amyloid; Down's syndrome

  Query Match 100.0%; Score 52; DB 2; Length 82;
  Best Local Similarity 100.0%; Pred. No. 0.0047;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQYTSIHHG 9
      |||||||||
Db 71 KQYTSIHHG 79

RESULT 2
A49795
Alzheimer's disease amyloid beta protein precursor - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A49795
R;Podlisny, M.B.; Tolan, D.R.; Selkoe, D.J.
Am. J. Pathol. 138, 1423-1435, 1991
A;Title: Homology of the amyloid beta protein precursor in monkey and human supports a primate model for beta amyloidosis in Alzheimer's disease.
A;Reference number: A49795; MUID:91273117; PMID:1905108
A;Accession: A49795
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-695 <POD>
A;Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing

  Query Match 100.0%; Score 52; DB 1; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.049;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQYTSIHHG 9
      |||||||||
Db 651 KQYTSIHHG 659

RESULT 3
A27485
Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse
N;Alternate names: proteinase nexin II
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999
C;Accession: A27485; S19727; I49485
R;Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sakaki, Y.
Biochem. Biophys. Res. Commun. 149, 665-671, 1987
A;Title: Complementary DNA for the mouse homolog of the human amyloid beta protein precursor.
A;Reference number: A27485; MUID:88106489; PMID:3322280
A;Accession: A27485
A;Molecule type: mRNA
A;Residues: 1-695 <YAM>
A;Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085
A;Experimental source: brain
R;de Strooper, B.; van Leuven, F.; van den Berghe, H.
Biochim. Biophys. Acta 1129, 141-143, 1991
A;Title: The amyloid beta protein precursor or proteinase nexin II from mouse is closer related to its human homolog than previously reported.
A;Reference number: S19727; MUID:92096458; PMID:1756177
A;Accession: S19727
A;Molecule type: mRNA
A;Residues: 1-210,'G',212-220,'S',222-396,'A',398-402,'T',404-448,'A',450-695 <STR>
A;Cross-references: EMBL:X59379
R;Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sakaki, Y.
Gene 112, 189-195, 1992
A;Title: Positive and negative regulatory elements for the expression of the Alzheimer's disease amyloid precursor-encoding gene in mouse.
A;Reference number: I49485; MUID:92209998; PMID:1555768
A;Accession: I49485
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-19 <RES>
A;Cross-references: GB:D10603; NID:g220328; PIDN:BAA01456.1; PID:g220329
C;Genetics:
A;Map position: 16C3
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid; transmembrane protein

  Query Match 100.0%; Score 52; DB 2; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.049;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQYTSIHHG 9
      |||||||||
Db 651 KQYTSIHHG 659

RESULT 4
S00550
Alzheimer's disease amyloid beta protein precursor - rat
N;Alternate names: beta-A4 amyloid protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C;Accession: S00550; A41245; A39820; S46251
R;Shivers, B.D.; Hilbich, C.; Multhaup, G.; Salbaum, M.; Beyreuther, K.; Seeburg, P.H.
EMBO J. 7, 1365-1370, 1988
A;Title: Alzheimer's disease amyloidogenic glycoprotein: expression pattern in rat brain suggests a role in cell contact.
A;Reference number: S00550; MUID:88312583; PMID:2900758
A;Accession: S00550

A;Molecule type: mRNA
A;Residues: 1-695 <SHI>
A;Cross-references: EMBL:X07648; NID:g55616; PIDN:CAA30488.1; PID:g55617
R;Schubert, D.; Schroeder, R.; LaCorbiere, M.; Saitoh, T.; Cole, G.
Science 241, 223-226, 1988
A;Title: Amyloid beta protein precursor is possibly a heparan sulfate proteoglycan core protein.
A;Reference number: A41245; MUID:88264430; PMID:2968652
A;Accession: A41245
A;Molecule type: protein
A;Residues: 18-37,'X',39-40,'X',42-44 <SCH>
A;Note: evidence for heparan sulfate attachment
R;Hesse, L.; Beher, D.; Masters, C.L.; Multhaup, G.
FEBS Lett. 349, 109-116, 1994
A;Title: The beta-A4 amyloid precursor protein binding to copper.
A;Reference number: S46251; MUID:94320627; PMID:7913895
A;Contents: annotation; copper binding sites
A;Note: rat peptides were isolated but not sequenced
R;Potempska, A.; Styles, J.; Mehta, P.; Kim, K.S.; Miller, D.L.
J. Biol. Chem. 266, 8464-8469, 1991
A;Title: Purification and tissue level of the beta-amyloid peptide precursor of rat brain.
A;Reference number: A39820; MUID:91217087; PMID:1673681
A;Accession: A39820
A;Status: preliminary
A;Molecule type: protein
A;Residues: 18-32 <POT>
A;Experimental source: brain
C;Comment: Deposition of amyloid protein as neurofibrillary tangles and/or plaques is characteristic of both Alzheimer's disease and Down's syndrome.
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid; glycoprotein; transmembrane protein
F;625-648/Domain: transmembrane #status predicted <TMM>

```
  Query Match             100.0%;  Score 52;  DB 2;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 0.049;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        651 KQYTSIHHG 659
```

RESULT 5
QRHUA4
Alzheimer's disease amyloid beta protein precursor [validated] - human
N;Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor XIa inhibitor; proteinase nexin II (PN-II)
N;Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascular form; amyloid protein precursor splice form APP(695); amyloid protein precursor splice form APP(751); amyloid protein precursor splice form APP(770)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 28-Jul-1995 #text_change 15-Sep-2000
C;Accession: S02260; S05194; A32277; A33260; A35486; I39452; I39451; I39453; I59562; A44017; B44017; A03134; A29030; A47584; A47585; S02638; S00707; S00925; A38949; A30320; B30320; C30320; A31087; A24668; A28583; A29302; A60805; JL0038; S06121; A60355; A59011; A38384; S29076; S38252; S32539; S48148; S48692; S51186; S51185; S51184; S51183; A54238; I58075; I52250; S09010; S10737; S24127; S43644
R;Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.; Beyreuther, K.; Mueller-Hill, B.
Nucleic Acids Res. 17, 517-522, 1989
A;Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encoded by 16 exons.
A;Reference number: S02260; MUID:89128427; PMID:2783775
A;Accession: S02260
A;Molecule type: DNA
A;Residues: 1-288,'V',365-770 <LEM1>
A;Cross-references: EMBL:X13466
A;Note: alternative splice form APP(695)
R;Lemaire, H.G.
submitted to the EMBL Data Library, November 1988
A;Reference number: S05194
A;Accession: S05194
A;Molecule type: DNA
A;Residues: 1-14,'VW',17-288,'V',365-770 <LEM2>
A;Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360
A;Note: alternative splice form APP(695)
R;La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.
Biochem. Biophys. Res. Commun. 159, 297-304, 1989
A;Title: Characterization of the 5'-end region and the first two exons of the beta-protein precursor gene.
A;Reference number: A32277; MUID:89165870; PMID:2538123
A;Accession: A32277
A;Molecule type: DNA
A;Residues: 1-75 <LAF>
A;Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AAC13654.1; PID:g516074
R;Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.
Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989
A;Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows similarity to soybean trypsin inhibitor.
A;Reference number: A33260; MUID:89392030; PMID:2675837
A;Accession: A33260
A;Molecule type: DNA
A;Residues: 656-737 <JOH>
A;Cross-references: GB:M29270; NID:g178863; PIDN:AAA51768.1; PID:g178865
R;Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B.
Biochem. Biophys. Res. Commun. 170, 301-307, 1990
A;Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of hereditary cerebral hemorrhage, Dutch type: DNA and protein diagnostic assays.
A;Reference number: A35486; MUID:90321244; PMID:2196878
A;Accession: A35486
A;Molecule type: DNA
A;Residues: 672-710 <PRE1>
A;Note: 693-Gln was found in DNA isolated from HCHWA-D patients
R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 87, 257-263, 1990
A;Title: Genomic organization of the human amyloid beta-protein precursor gene.
A;Reference number: I39451; MUID:90236318; PMID:2110105
A;Accession: I39452

A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-770 <YOS1>
A;Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616
A;Accession: I39451
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-530,'QWLMPVIPAFWEAKVGR' <YOS2>
A;Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615
R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 102, 291-292, 1991
A;Reference number: A59020; MUID:91340168; PMID:1908403
A;Contents: annotation; erratum
A;Note: revised physical map for reference I39451
R;Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Duinen, S.G.; Bots, G.T.; Luyendijk, W.; Frangione, B.
Science 248, 1124-1126, 1990
A;Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type.
A;Reference number: I39453; MUID:90260663; PMID:2111584
A;Accession: I39453
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 656-737 <LEV>
A;Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620
A;Note: a mutation with 693-Gln is presented
R;Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A;Title: A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease.
A;Reference number: I59562; MUID:92022553; PMID:1925564
A;Accession: I59562
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 689-716,'F',718-737 <MUR>
A;Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721
R;Kamino, K.; Orr, H.T.; Payami, H.; Wijsman, E.M.; Alonso, M.E.; Pulst, S.M.; Anderson, L.; O'dahl, S.; Nemens, E.; White, J.A.; Sadovnick, A.D.; Ball, M.J.; Kaye, J.; Warren, A.; McInnis, M.; Antonarakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin, G.M.; Bird, T.D.; Schellenberg, G.D.
Am. J. Hum. Genet. 51, 998-1014, 1992
A;Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for the APP gene region.
A;Reference number: A44017; MUID:93035397; PMID:1415269
A;Accession: A44017
A;Molecule type: DNA
A;Residues: 687-692,'G',694-718 <KAM1>
A;Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378
A;Experimental source: familial Alzheimer disease family SB
A;Note: sequence extracted from NCBI backbone (NCBIP:115374)
A;Accession: B44017
A;Molecule type: DNA
A;Residues: 687-718 <KAM2>
A;Cross-references: GB:S45136; NID:g257379; PIDN:AAB23646.1; PID:g257380
A;Experimental source: familial Alzheimer disease family LIT
A;Note: sequence extracted from NCBI backbone (NCBIP:115376)
A;Note: this sequence has a silent mutation
R;Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.H.; Multhaup, G.; Beyreuther, K.; Muller-Hill, B.
Nature 325, 733-736, 1987
A;Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor.
A;Reference number: A03134; MUID:87144572; PMID:2881207
A;Accession: A03134
A;Molecule type: mRNA
A;Residues: 1-288,'V',365-770 <KAN>
A;Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
A;Note: alternative splice form APP(695)
R;Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A;Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascular and the neuritic plaque amyloid peptides.
A;Reference number: A29030; MUID:87231971; PMID:3035574
A;Accession: A29030
A;Molecule type: mRNA
A;Residues: 284-288,'V',365-646,'E',648-770 <ROB>
A;Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540
A;Note: the authors translated the codon GAG for residue 647 as Asp
R;Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffiotti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987
A;Title: Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease.
A;Reference number: A47584; MUID:87120328; PMID:3810169
A;Accession: A47584
A;Molecule type: mRNA
A;Residues: 674-756,'S',758-770 <GOL>
A;Cross-references: GB:M15533; NID:g178706; PIDN:AAA35540.1; PID:g178707
A;Experimental source: brain
R;Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van Keuren, M.L.; Patterson, D.; Pagan, S.; Kurnit, D.M.; Neve, R.L.
Science 235, 880-884, 1987
A;Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus.
A;Reference number: A47585; MUID:87120329; PMID:2949367
A;Accession: A47585
A;Molecule type: mRNA
A;Residues: 674-703 <TAN1>
A;Cross-references: GB:M15532; NID:g177957; PIDN:AAA51564.1; PID:g177958
R;Dyrks, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mueller-Hill, B.; Masters, C.L.; Beyreuther, K.
EMBO J. 7, 949-957, 1988
A;Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 precursor of Alzheimer's disease.
A;Reference number: S02638; MUID:88296437; PMID:2900137
A;Accession: S02638
A;Molecule type: mRNA
A;Residues: 672-678 <DYR>
R;Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; Neve, R.L.
Nature 331, 528-530, 1988

A;Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease.
A;Reference number: S00707; MUID:88122640; PMID:2893290
A;Accession: S00707
A;Molecule type: mRNA
A;Residues: 286-344,'I',365-366 <TAN2>
A;Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g929612
A;Experimental source: promyelocytic leukemia cell line HL60
A;Note: alternative splice form APP(751)
R;Ponte, P.; Gonzalez-DeWhitt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.; Davis, K.; Wallace, W.; Lieberburg, I.; Fuller, F.; Cordell, B.
Nature 331, 525-527, 1988
A;Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors.
A;Reference number: S00925; MUID:88122639; PMID:2893289
A;Accession: S00925
A;Molecule type: mRNA
A;Residues: 1-344,'I',365-770 <PO2>
A;Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721
A;Note: alternative splice form APP(751)
R;Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
Nature 331, 530-532, 1988
A;Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity.
A;Reference number: A38949; MUID:88122641; PMID:2893291
A;Accession: A38949
A;Molecule type: mRNA
A;Residues: 287-367 <KIT>
A;Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g929611
A;Experimental source: glioblastoma cell line
A;Note: alternative splice form APP(770)
R;Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ashton, R.A.; Macq, A.F.; Maloteaux, J.M.; Blume, A.J.; Octave, J.N.
Brain Res. Mol. Brain Res. 4, 121-131, 1988
A;Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three patients with sporadic Alzheimer's disease.
A;Reference number: A30320
A;Accession: A30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 284-288,'V',365-770 <VIT1>
A;Accession: B30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 122-288,'V',365-770 <VIT2>
A;Accession: C30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 606-770 <VIT3>
R;Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.; Marotta, C.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
A;Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease brain: coding and noncoding regions of the fetal precursor mRNA are expressed in the cortex.
A;Reference number: A31087; MUID:88124954; PMID:2893379
A;Accession: A31087
A;Molecule type: mRNA
A;Residues: 507-770 <ZAI>
A;Cross-references: GB:M18734; NID:g178572; PIDN:AAA51726.1; PID:g178573
A;Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 603 as Val, GTG for residue 604 as Glu, GAG for residue 605 as Leu, CTT for residue 607 as Pro, CCC for residue 608 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 655 as Ser
A;Note: the cited Genbank accession number, J03594, is not in release 101.0
R;Masters, C.L.; Multhaup, G.; Simms, G.; Pottgiesser, J.; Martins, R.N.; Beyreuther, K.

Query Match 100.0%; Score 52; DB 1; Length 770;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQYTSIHHG 9
              |||||||||
Db        726 KQYTSIHHG 734
```

Search completed: October 4, 2004, 18:58:23
Job time : 20.3043 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:44:39 ; Search time 7.43478 Seconds
(without alignments)
63.032 Million cell updates/sec

Title: US-10-048-209-2
Perfect score: 52
Sequence: 1 KQYTSIHHG 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	52	100.0	737	1	A4_FUGRU	O93279 fugu rubrip
2	52	100.0	751	1	A4_SAISC	Q95241 s amyloid b
3	52	100.0	770	1	A4_CAVPO	Q60495 c amyloid b
4	52	100.0	770	1	A4_HUMAN	P05067 h amyloid b
5	52	100.0	770	1	A4_MACFA	P53601 m amyloid b
6	52	100.0	770	1	A4_MOUSE	P12023 m amyloid b
7	52	100.0	770	1	A4_PIG	P79307 s amyloid b
8	52	100.0	770	1	A4_RAT	P08592 r amyloid b
9	52	100.0	780	1	A4_TETFL	O73683 tetraodon f
10	39	75.0	452	1	NU4M_LUMTE	Q34949 lumbricus t
11	38	73.1	488	1	YGM9_YEAST	Q01163 saccharomyc
12	35	67.3	197	1	PSD9_CAEEL	Q10920 caenorhabdi
13	35	67.3	443	1	GAT3_HUMAN	P23771 homo sapien
14	35	67.3	443	1	GAT3_MOUSE	P23772 mus musculu
15	35	67.3	739	1	PURL_LISIN	Q92an9 listeria in
16	35	67.3	739	1	PURL_LISMO	Q8y6c1 listeria mo
17	35	67.3	1268	1	YRD6_CAEEL	Q09575 caenorhabdi
18	34	65.4	1406	1	TOP1_CANGA	O93794 candida gla
19	33	63.5	117	1	HEMU_LINUN	P22765 lingula ung
20	33	63.5	142	1	Y4KQ_RHISN	P55535 rhizobium s
21	33	63.5	157	1	SSRP_CHLTE	Q8kew8 chlorobium
22	33	63.5	339	1	TF3A_XENBO	P17842 xenopus bor
23	33	63.5	361	1	SERC_BACHD	Q9kdm4 bacillus ha
24	33	63.5	366	1	TF3A_XENLA	P03001 xenopus lae
25	33	63.5	379	1	ISPH_SYNY3	Q55643 synechocyst
26	33	63.5	519	1	LEU1_CANBF	Q7vqj6 candidatus
27	33	63.5	530	1	UDBC_RAT	P36511 rattus norv
28	33	63.5	530	1	UDBE_RABIT	P36513 oryctolagus
29	33	63.5	645	1	ACS2_PSEAE	Q9hv66 pseudomonas
30	33	63.5	683	1	CV19_HUMAN	Q13769 homo sapien
31	33	63.5	805	1	DF19_CAEEL	Q09555 caenorhabdi
32	33	63.5	4344	1	DYHC_EMENI	P45444 emericella
33	32	61.5	121	1	OMP7_STAAU	P21223 staphylococ
34	32	61.5	239	1	DCUR_ECOL6	P59338 escherichia
35	32	61.5	239	1	DCUR_ECOLI	P39271 escherichia
36	32	61.5	239	1	DCUR_SHIFL	P59339 shigella fl
37	32	61.5	249	1	CREB_CHLVR	P51984 chlorohydra
38	32	61.5	284	1	YMX7_YEAST	Q04299 saccharomyc
39	32	61.5	326	1	ZN73_HUMAN	O43830 homo sapien
40	32	61.5	340	1	LPXD_WIGBR	Q8d2h2 wiggleswort
41	32	61.5	352	1	GBA1_COPCO	P30675 coprinus co
42	32	61.5	510	1	CF2_DROME	P20385 drosophila
43	32	61.5	579	1	YC12_KLEPN	Q48458 klebsiella
44	32	61.5	580	1	SYD_MYCPE	Q8ewb7 mycoplasma
45	32	61.5	717	1	YCCS_ECOLI	P75870 escherichia

ALIGNMENTS

RESULT 1
A4_FUGRU
ID A4_FUGRU STANDARD; PRT; 737 AA.
AC O93279;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog precursor [Contains:
DE Beta-amyloid protein (Beta-APP) (A-beta)].
GN APP.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetradontoidea; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98252138; PubMed=9599080;
RA Villard L., Tassone F., Crnogorac-Jurcevic T., Clancy K., Gardiner K.;
RT "Analysis of pufferfish homologues of the AT-rich human APP gene.";
RL Gene 210:17-24(1998).
CC -!- FUNCTION: Functional neuronal receptor which couples to
CC intracellular signaling pathway through the GTP-binding protein .

```
CC       G(O) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF090120; AAD13392.1; -.
DR   HSSP; P05067; 1HZ3.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; FALSE_NEG.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
KW   Serine protease inhibitor.
FT   SIGNAL        1    18       POTENTIAL.
FT   CHAIN        19   737       ALZHEIMER'S DISEASE AMYLOID A4
FT                               PROTEIN HOMOLOG.
FT   CHAIN       639   681       BETA-AMYLOID PROTEIN (POTENTIAL).
FT   DOMAIN       19   668       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    669   689       POTENTIAL.
FT   DOMAIN      690   737       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      286   344       BPTI/KUNITZ INHIBITOR.
FT   SITE        726   729       CLATHRIN-BINDING (BY SIMILARITY).
FT   ACT_SITE    300   301       REACTIVE BOND.
FT   DISULFID    290   340       BY SIMILARITY.
FT   DISULFID    299   323       BY SIMILARITY.
FT   DISULFID    315   336       BY SIMILARITY.
FT   CARBOHYD    522   522       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   737 AA;  82856 MW;  6FAD01E2E3B2B7E2 CRC64;

  Query Match             100.0%;  Score 52;  DB 1;  Length 737;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        693 KQYTSIHHG 701
```

```
RESULT 2
A4_SAISC
ID   A4_SAISC       STANDARD;      PRT;   751 AA.
AC   Q95241;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid
DE   protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha); Soluble
DE   APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42);
DE   Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-
DE   CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
GN   APP.
OS   Saimiri sciureus (Common squirrel monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX   NCBI_TaxID=9521;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Liver;
RX   MEDLINE=96108492; PubMed=8532114;
RA   Levy E., Amorim A., Frangione B., Walker L.C.;
RT   "Beta-amyloid precursor protein gene in squirrel monkeys with
RT   cerebral amyloid angiopathy.";
RL   Neurobiol. Aging 16:805-808(1995).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction. In vitro, copper-metallated APP induces neuronal
CC       death directly or is potentiated through Cu(II)-mediated low-
CC       density lipoprotein oxidation (By similarity). Can regulate
CC       neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
```

```
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=APP770;
CC         IsoId=Q95241-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=Q95241-2; Sequence=Not described;
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-720 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S81024; AAD14347.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.

DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Amyloid; Alternative splicing.
FT   SIGNAL       1     17       BY SIMILARITY.
FT   CHAIN       18    751       A4 PROTEIN.
FT   CHAIN       18    668       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN       18    652       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN      653    751       C99 (POTENTIAL).
FT   CHAIN      653    694       BETA-AMYLOID PROTEIN 42 (POTENTIAL).
FT   CHAIN      653    692       BETA-AMYLOID PROTEIN 40 (POTENTIAL).
FT   CHAIN      669    751       C83 (POTENTIAL).
FT   CHAIN      669    694       P3(42) (POTENTIAL).
FT   CHAIN      669    692       P3(40) (POTENTIAL).
FT   CHAIN      693    751       GAMMA-CTF(59) (POTENTIAL).
FT   CHAIN      695    751       GAMMA-CTF(57) (POTENTIAL).
FT   CHAIN      702    751       GAMMA-CTF(50) (POTENTIAL).
FT   CHAIN      721    751       C31 (POTENTIAL).
FT   DOMAIN      18    680       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   681    704       POTENTIAL.
FT   DOMAIN     705    751       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      96    110       HEPARIN-BINDING (BY SIMILARITY).
```

```
FT   DOMAIN      181    188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN      291    341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN      316    344       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      363    428       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      504    521       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN      713    732       INTERACTION WITH G(O)-ALPHA
FT                                (BY SIMILARITY).
FT   DOMAIN      230    260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      274    280       POLY-THR.
FT   SITE        144    144       REQUIRED FOR COPPER(II) REDUCTION
FT                                (BY SIMILARITY).
FT   ACT_SITE    301    302       REACTIVE BOND.
FT   SITE        652    653       CLEAVAGE (BY BETA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        653    654       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE        668    669       CLEAVAGE (BY ALPHA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        685    685       INVOLVED IN FREE RADICAL PROPAGATION
FT                                (BY SIMILARITY).
FT   SITE        687    687       INVOLVED IN OXIDATIVE REACTIONS
FT                                (BY SIMILARITY).
FT   SITE        692    693       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                                (BY SIMILARITY).
FT   SITE        694    695       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                                (BY SIMILARITY).
FT   SITE        701    702       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT                                (BY SIMILARITY).
FT   SITE        705    715       BASOLATERAL SORTING SIGNAL
FT                                (BY SIMILARITY).
FT   SITE        720    721       CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)
FT                                (BY SIMILARITY).
FT   SITE        738    741       ENDOCYTOSIS SIGNAL.
FT   SITE        740    743       NPXY MOTIF.

  Query Match             100.0%;  Score 52;  DB 1;  Length 751;
  Best Local Similarity   100.0%;  Pred. No. 0.016;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        707 KQYTSIHHG 715


RESULT 3
A4_CAVPO
ID   A4_CAVPO        STANDARD;      PRT;   770 AA.
AC   Q60495; Q60496;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); CTF-alpha; CTF-beta; Beta-amyloid
DE   protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); P3(42);
DE   P3(40); CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-
DE   CTF(57) (Gamma-secretase C-terminal fragment 57); C31].
GN   APP.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain, and Liver;
RX   MEDLINE=97236426; PubMed=9116031;
RA   Beck M., Mueller D., Bigl V.;
RT   "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
RT   alternative splicing.";
RL   Biochim. Biophys. Acta 1351:17-21(1997).
RN   [2]
RP   INTERACTION OF BETA-APP40 WITH APOE.
RX   MEDLINE=98007700; PubMed=9349544;
RA   Martel C.L., Mackic J.B., Matsubara E., Governale S., Miguel C.,
RA   Miao W., McComb J.G., Frangione B., Ghiso J., Zlokovic B.V.;
RT   "Isoform-specific effects of apolipoproteins E2, E3, and E4 on
RT   cerebral capillary sequestration and blood-brain barrier transport of
RT   circulating Alzheimer's amyloid beta.";
RL   J. Neurochem. 69:1995-2004(1997).
RN   [3]
RP   PROCESSING.
RX   MEDLINE=20084499; PubMed=10619481;
RA   Beck M., Brueckner M.K., Holzer M., Kaap S., Pannicke T., Arendt T.,
RA   Bigl V.;
RT   "Guinea-pig primary cell cultures provide a model to study expression
RT   and amyloidogenic processing of endogenous amyloid precursor
RT   protein.";
RL   Neuroscience 95:243-254(2000).
RN   [4]
RP   GAMMA-SECRETASE PROCESSING.
RX   MEDLINE=20576391; PubMed=11035007;
RA   Pinnix I., Musunuru U., Tun H., Sridharan A., Golde T., Eckman C.,
RA   Ziani-Cherif C., Onstead L., Sambamurti K.;
RT   "A novel gamma -secretase assay based on detection of the putative
RT   C-terminal fragment-gamma of amyloid beta protein precursor.";
RL   J. Biol. Chem. 276:481-487(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction (By similarity). In vitro, copper-metallated APP
CC       induces neuronal death directly or is potentiated through Cu(II)-
CC       mediated low-density lipoprotein oxidation (By similarity). Can
CC       regulate neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
```

```
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Beta-amyloid peptides bind to lipoproteins
CC       and apoliproteins E and J in the CSF and to HDL particles in
CC       plasma, inhibiting metal-catalyzed oxidation of lipoproteins.
CC   -!- FUNCTION: Appicans elicit adhesion of neural cells to the
CC       extracellular matrix and may regulate neurite outgrowth in the
CC       brain (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains), APPBP2 (via BaSS) and DDB1 (By similarity).
CC       Associates with microtubules in the presence of ATP and in a
CC       kinesin-dependent manner (By similarity). Soluble Abeta40 binds
CC       all three isoforms of APOE, in vitro and in vivo. When lipidated,
CC       ApoE3 appears to be the preferred amyloid binding isoform, while
CC       the apoE4 isoform-beta-APP40 complex is capable of being
CC       transported across the blood-brain barrier.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated pits
CC       (By similarity). During maturation, the immature APP (N-
CC       glycosylated in the endoplasmic reticulum) moves to the Golgi
CC       complex where complete maturation occurs (O-glycosylated and
CC       sulfated) (By similarity). After alpha-secretase cleavage, soluble
CC       APP is released into the extracellular space and the C-terminal is
CC       internalized to endosomes and lysosomes (By similarity). Some APP
CC       accumulates in secretory transport vesicles leaving the late Golgi
CC       compartment and returns to the cell surface (By similarity). APP
CC       sorts to the basolateral surface in epithelial cells (By
CC       similatity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms, missing exons 7,8 and 15, seem to
CC         exist. The L-isoforms, missing exon 15, are referred to as
CC         appicans;
CC       Name=APP770;
CC         IscId=Q60495-1; Sequence=Displayed;
CC       Name=APP695;
CC         IscId=Q60495-2; Sequence=VSP_007221, VSP_007222;
CC   -!- TISSUE SPECIFICITY: Isoform APP695 is the major isoform found in
CC       brain. The longer isoforms containing the BPTI domain are
CC       predominantly expressed in peripheral organs such as muscle and
CC       liver.
CC   -!- INDUCTION: Increased levels during neuronal differentiation.
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells.
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue (By similarity). The NPXY site is also involved in
CC       clathrin-mediated endocytosis.
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       CTF-alpha and CTF-beta. Subsequent processing of CTF-alpha by
CC       gamma-secretase yields P3 peptides. This is the major secretory
CC       pathway and is nonamyloidogenic. Alternatively,
CC       presenilin/nicastrin-mediated gamma-secretase processing of CTF-
CC       beta releases the amyloid beta proteins, amyloid-beta 40 (Abeta40)
CC       and amyloid-beta 42 (Abeta42), major components of amyloid
CC       plaques, and the corresponding cytotoxic C-terminal fragments
CC       (CTFs).
CC   -!- PTM: Proteolytically cleaved by caspase-3 during neuronal
CC       apoptosis (By similarity).
CC   -!- PTM: N- and O-glycosylated. O-linkage of chondroitin sulfate to
CC       the L-APP isoforms produces the APP proteoglycan core proteins,
CC       the appicans (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific (By similarity).
CC       Phosphorylation can affect APP processing, neuronal
CC       differentiation and interaction with other proteins.
CC   -!- PTM: Extracellular binding and reduction of copper, results in a
CC       corresponding oxidation of Cys-144 and Cys-158, and the formation
CC       of a disulfide bond (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates.
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X97631; CAA66230.1; -.
DR   EMBL; X99198; CAA67589.1; -.
DR   HSSP; P05067; 1BA4.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
```

```
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Alternative splicing; Amyloid.
FT   SIGNAL        1     17       BY SIMILARITY.
FT   CHAIN        18    770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN        18    687       SOLUBLE APP-ALPHA (BY SIMILARITY).
FT   CHAIN        18    671       SOLUBLE APP-BETA (BY SIMILARITY).
FT   CHAIN       672    770       CTF-ALPHA (BY SIMILARITY).
FT   CHAIN       672    713       BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
FT   CHAIN       672    711       BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
FT   CHAIN       688    770       CTF-BETA (BY SIMILARITY).
FT   CHAIN       688    713       P3(42) (BY SIMILARITY).
FT   CHAIN       688    711       P3(40) (BY SIMILARITY).
FT   CHAIN       712    770       GAMMA-CTF(59) (BY SIMILARITY).
FT   CHAIN       714    770       GAMMA-CTF(57) (BY SIMILARITY).

  Query Match             100.0%;  Score 52;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        726 KQYTSIHHG 734


RESULT 4
A4_HUMAN
ID   A4_HUMAN        STANDARD;      PRT;   770 AA.
AC   P05067; P09000; P78438; Q13764; Q13778; Q13793; Q16011; Q16014;
AC   Q16019; Q16020; Q9BT38; Q9UCA9; Q9UCB6; Q9UCC8; Q9UCD1; Q9UQ58;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein) (Cerebral vascular amyloid peptide) (CVAP) (Protease
DE   nexin-II) (PN-II) (APPI) (PreA4) [Contains: Soluble APP-alpha (S-APP-
DE   alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42
DE   (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42);
DE   P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59)
DE   (Amyloid intracellular domain 59) (AID(59)); Gamma-CTF(57) (Gamma-
DE   secretase C-terminal fragment 57) (Amyloid intracellular domain 57)
DE   (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50)
DE   (Amyloid intracellular domain 50) (AID(50)); C31].
GN   APP OR A4 OR AD1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=87144572; PubMed=2881207;
RA   Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,
RA   Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;
RT   "The precursor of Alzheimer's disease amyloid A4 protein resembles a
RT   cell-surface receptor.";
RL   Nature 325:733-736(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM APP751).
RC   TISSUE=Brain;
RX   MEDLINE=88122639; PubMed=2893289;
RA   Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D.,
RA   Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,
RA   Cordell B.;
RT   "A new A4 amyloid mRNA contains a domain homologous to serine
RT   proteinase inhibitors.";
RL   Nature 331:525-527(1988).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RX   MEDLINE=89128427; PubMed=2783775;
RA   Lemaire H.-G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,
RA   Unterbeck A., Beyreuther K., Mueller-Hill B.;
RT   "The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid
RT   is encoded by 16 exons.";
RL   Nucleic Acids Res. 17:517-522(1989).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM APP770).
RX   MEDLINE=90236318; PubMed=2110105;
RA   Yoshikai S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
RT   "Genomic organization of the human amyloid beta-protein precursor
RT   gene.";
RL   Gene 87:257-263(1990).
RN   [5]
RP   ERRATUM, AND REVISIONS.
RA   Yoshikai S.-I., Sasaki H., Doh-ura K., Furuya H., Sakaki Y.;
RL   Gene 102:291-292(1991).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM L-APP733).
RC   TISSUE=Leukocyte;
RX   MEDLINE=92268136; PubMed=1587857;
RA   Koenig G., Moenning U., Czech C., Prior R., Banati R.,
RA   Schreiter-Gasser U., Bauer J., Masters C.L., Beyreuther K.;
RT   "Identification and differential expression of a novel alternative
RT   splice isoform of the beta A4 amyloid precursor protein (APP) mRNA in
RT   leukocytes and brain microglial cells.";
RL   J. Biol. Chem. 267:10804-10809(1992).
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM APP770).
RX   MEDLINE=97263807; PubMed=9108164;
RA   Hattori M., Tsukahara F., Furuhata Y., Tanahashi H., Hirose M.,
RA   Saito M., Tsukuni S., Sakaki Y.;
RT   "A novel method for making nested deletions and its application for
RT   sequencing of a 300 kb region of human APP locus.";
RL   Nucleic Acids Res. 25:1802-1808(1997).
RN   [8]
RP   SEQUENCE FROM N.A. (ISOFORM APP639).
RC   TISSUE=Brain;
RX   MEDLINE=22744650; PubMed=12859342;
RA   Tang K., Wang C., Shen C., Sheng S., Ravid R., Jing N.;
RT   "Identification of a novel alternative splicing isoform of human
RT   amyloid precursor protein gene, APP639.";
RL   Eur. J. Neurosci. 18:102-108(2003).
```

RN [9]
RP SEQUENCE FROM N.A. (ISOFORM APP305).
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [10]
RP SEQUENCE OF 1-10 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016647; PubMed=3140222;
RA Schon E.A., Mita S., Sadlock J., Herbert J.;
RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
RT encodes a 95-kDa polypeptide.";
RL Nucleic Acids Res. 16:9351-9351(1988).
RN [11]
RP ERRATUM, AND REVISIONS.
RA Mita S., Sadlock J., Herbert J., Schon E.A.;
RL Nucleic Acids Res. 16:11402-11402(1988).
RN [12]
RP SEQUENCE OF 1-75 FROM N.A.
RX MEDLINE=89165870; PubMed=2538123;
RA La Fauci G., Lahiri D.K., Salton S.R., Robakis N.K.;
RT "Characterization of the 5'-end region and the first two exons of the
RT beta-protein precursor gene.";
RL Biochem. Biophys. Res. Commun. 159:297-304(1989).
RN [13]
RP SEQUENCE OF 18-50.
RC TISSUE=Fibroblast;
RX MEDLINE=87250462; PubMed=3597385;
RA van Nostrand W.E., Cunningham D.D.;
RT "Purification of protease nexin II from human fibroblasts.";
RL J. Biol. Chem. 262:8508-8514(1987).
RN [14]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP751).
RC TISSUE=Brain;
RX MEDLINE=89346754; PubMed=2569763;
RA de Sauvage F., Octave J.N.;
RT "A novel mRNA of the A4 amyloid precursor gene coding for a possibly
RT secreted protein.";
RL Science 245:651-653(1989).
RN [15]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP695).
RC TISSUE=Brain;
RX MEDLINE=87231971; PubMed=3035574;
RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
RT "Molecular cloning and characterization of a cDNA encoding the
RT cerebrovascular and the neuritic plaque amyloid peptides.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
RN [16]
RP SEQUENCE OF 286-366 FROM N.A.
RX MEDLINE=88122640; PubMed=2893290;
RA Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,
RA Gusella J.F., Neve R.L.;
RT "Protease inhibitor domain encoded by an amyloid protein precursor
RT mRNA associated with Alzheimer's disease.";
RL Nature 331:528-530(1988).
RN [17]
RP SEQUENCE OF 287-367 FROM N.A.
RX MEDLINE=88122641; PubMed=2893291;
RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
RT "Novel precursor of Alzheimer's disease amyloid protein shows
RT protease inhibitory activity.";
RL Nature 331:530-532(1988).
RN [18]
RP SEQUENCE OF 507-770 FROM N.A.
RC TISSUE=Brain cortex;
RX MEDLINE=88124954; PubMed=2893379;
RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
RA Marotta C.A.;
RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
RT disease brain: coding and noncoding regions of the fetal precursor
RT mRNA are expressed in the cortex.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
RN [19]
RP SEQUENCE OF 523-555, AND COLLAGEN-BINDING DOMAIN.
RX MEDLINE=96139497; PubMed=8576160;
RA Beher D., Hesse L., Masters C.L., Multhaup G.;
RT "Regulation of amyloid protein precursor (APP) binding to collagen and
RT mapping of the binding sites on APP and collagen type I.";
RL J. Biol. Chem. 271:1613-1620(1996).
RN [20]
RP SEQUENCE OF 655-737 FROM N.A., AND VARIANTS AD PHE-717; AD ILE-717
RP AND AD GLY-717.
RX MEDLINE=93236601; PubMed=8476439;
RA Denman R.B., Rosenzcwaig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
RN [21]
RP SEQUENCE OF 656-737 FROM N.A.
RX MEDLINE=89392030; PubMed=2675837;
RA Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
RA Little S.P.;
RT "Alzheimer's disease amyloid peptide is encoded by two exons and shows
RT similarity to soybean trypsin inhibitor.";
RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).

```
RN   [22]

  Query Match             100.0%;  Score 52;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        726 KQYTSIHHG 734


RESULT 5
A4_MACFA
ID   A4_MACFA       STANDARD;      PRT;   770 AA.
AC   P53601; Q95KN7;
DT   01-OCT-1996 (Rel. 34, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-
DE   APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);
DE   Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
GN   APP.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS APP695 AND APP770).
RC   TISSUE=Cerebellum;
RX   MEDLINE=91273117; PubMed=1905108;
RA   Podlisny M.B., Tolan D.R., Selkoe D.J.;
RT   "Homology of the amyloid beta protein precursor in monkey and human
RT   supports a primate model for beta amyloidosis in Alzheimer's
RT   disease.";
RL   Am. J. Pathol. 138:1423-1435(1991).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction. In vitro, copper-metallated APP induces neuronal
CC       death directly or is potentiated through Cu(II)-mediated low-
CC       density lipoprotein oxidation (By similarity). Can regulate
CC       neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=APP770;
CC         IsoId=P53601-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=P53601-2; Sequence=VSP_000010, VSP_000011;
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
```

```
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M58727; AAA36829.1; -.
DR   EMBL; M58726; AAA36828.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Alternative splicing; Amyloid.
FT   SIGNAL        1     17       BY SIMILARITY.
FT   CHAIN        18    770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN        18    687       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN        18    671       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN       672    770       C99 (POTENTIAL).
FT   CHAIN       672    713       BETA-AMYLOID PROTEIN 42 (POTENTIAL).
FT   CHAIN       672    711       BETA-AMYLOID PROTEIN 40 (POTENTIAL).
FT   CHAIN       688    770       C83 (POTENTIAL).
FT   CHAIN       688    713       P3(42) (POTENTIAL).
FT   CHAIN       688    711       P3(40) (POTENTIAL).
FT   CHAIN       712    770       GAMMA-CTF(59) (POTENTIAL).
FT   CHAIN       714    770       GAMMA-CTF(57) (POTENTIAL).
FT   CHAIN       721    770       GAMMA-CTF(50) (POTENTIAL).
FT   CHAIN       740    770       C31 (POTENTIAL).
FT   DOMAIN       18    699       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    700    723       POTENTIAL.
FT   DOMAIN      724    770       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       96    110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      181    188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN      291    341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN      391    423       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      491    522       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      523    540       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN      732    751       INTERACTION WITH G(O)-ALPHA
FT                                (BY SIMILARITY).
FT   DOMAIN      230    260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      274    280       POLY-THR.
FT   SITE        144    144       REQUIRED FOR COPPER(II) REDUCTION
FT                                (BY SIMILARITY).
FT   ACT_SITE    301    302       REACTIVE BOND (BY SIMILARITY).
FT   SITE        671    672       CLEAVAGE (BY BETA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        672    673       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE        687    688       CLEAVAGE (BY ALPHA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        704    704       IMPLICATED IN FREE RADICAL PROPAGATION
FT                                (BY SIMILARITY).
FT   SITE        706    706       INVOLVED IN OXIDATIVE REACTIONS
FT                                (BY SIMILARITY).
FT   SITE        711    712       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                                (BY SIMILARITY).
FT   SITE        713    714       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                                (BY SIMILARITY).
FT   SITE        720    721       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT                                (BY SIMILARITY).
FT   SITE        724    734       BASOLATERAL SORTING SIGNAL
FT                                (BY SIMILARITY).
FT   SITE        739    740       CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)

  Query Match             100.0%;  Score 52;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        726 KQYTSIHHG 734


RESULT 6
A4_MOUSE
ID   A4_MOUSE       STANDARD;      PRT;   770 AA.
AC   P12023; P97487; P97942; Q99K32;
DT   01-OCT-1989 (Rel. 12, Created)
```

```
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains:
DE   Soluble APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99
DE   (APP-C99); Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein
DE   40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase
DE   C-terminal fragment 59) (Amyloid intracellular domain 59) (AID(59))
DE   (APP-C59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57)
DE   (Amyloid intracellular domain 57) (AID(57)) (APP-C57); Gamma-CTF(50)
DE   (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain
DE   50) (AID(50)); C31].
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=88106489; PubMed=3322280;
RA   Yamada T., Sasaki H., Furuya H., Miyata T., Goto I., Sakaki Y.;
RT   "Complementary DNA for the mouse homolog of the human amyloid beta
RT   protein precursor.";
RL   Biochem. Biophys. Res. Commun. 149:665-671(1987).
RN   [2]
RP   REVISIONS.
RA   Yamada T.;
RL   Submitted (MAR-1988) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=92096458; PubMed=1756177;
RA   de Strooper B., van Leuven F., van den Berghe H.;
RT   "The amyloid beta protein precursor or proteinase nexin II from mouse
RT   is closer related to its human homolog than previously reported.";
RL   Biochim. Biophys. Acta 1129:141-143(1991).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   STRAIN=SAMP8; TISSUE=Hippocampus;
RX   MEDLINE=21130647; PubMed=11235921;
RA   Kumar V.B., Vyas K., Franko M., Choudhary V., Buddhiraju C.,
RA   Alvarez J., Morley J.E.;
RT   "Molecular cloning, expression, and regulation of hippocampal amyloid
RT   precursor protein of senescence accelerated mouse (SAMP8).";
RL   Biochem. Cell Biol. 79:57-67(2001).
RN   [5]
RP   SEQUENCE OF 1-19 FROM N.A.
RX   MEDLINE=92209998; PubMed=1555768;
RA   Izumi R., Yamada T., Yoshikai S.I., Sasaki H., Hattori M.,
RA   Sakai Y.;
RT   "Positive and negative regulatory elements for the expression of the
RT   Alzheimer's disease amyloid precursor-encoding gene in mouse.";
RL   Gene 112:189-195(1992).
RN   [6]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORM APP770).
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain, and Kidney;
RX   MEDLINE=89149813; PubMed=2493250;
RA   Yamada T., Sasaki H., Dohura K., Goto I., Sakaki Y.;
RT   "Structure and expression of the alternatively-spliced forms of mRNA
RT   for the mouse homolog of Alzheimer's disease amyloid beta protein
RT   precursor.";
RL   Biochem. Biophys. Res. Commun. 158:906-912(1989).
RN   [8]
RP   SEQUENCE OF 289-364 FROM N.A.
RC   STRAIN=CD-1; TISSUE=Placenta;
RX   MEDLINE=89345111; PubMed=2569710;
RA   Fukuchi K., Martin G.M., Deeb S.S.;
RT   "Sequence of the protease inhibitor domain of the A4 amyloid protein
RT   precursor of Mus domesticus.";
RL   Nucleic Acids Res. 17:5396-5396(1989).
RN   [9]
RP   SEQUENCE OF 656-737 FROM N.A.
RC   STRAIN=129/Sv;
RA   Wragg M.A., Busfield F., Duff K., Korenblat K., Capecchi M.,
RA   Loring J.F., Goate A.M.;
RT   "Introduction of six mutations into the mouse genome using 'Hit and
RT   Run' gene-targeting: introduction of familial Alzheimer's disease
RT   mutations into the mouse amyloid precursor protein gene and
RT   humanization of the A-beta fragment.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   TISSUE SPECIFICITY OF ALTERNATIVE SPLICED FORMS.
RX   MEDLINE=93287808; PubMed=8510506;
RA   Sola C., Mengod G., Ghetti B., Palacios J.M., Triarhou L.C.;
RT   "Regional distribution of the alternatively spliced isoforms of beta
RT   APP RNA transcript in the brain of normal, heterozygous and
RT   homozygous weaver mutant mice as revealed by in situ hybridization
RT   histochemistry.";
RL   Brain Res. Mol. Brain Res. 17:340-346(1993).
```

```
RN   [11]
RP   INTERACTION WITH KNS2.
RX   MEDLINE=21010507; PubMed=11144355;
RA   Kamal A., Stokin G.B., Yang Z., Xia C.-H., Goldstein L.S.;
RT   "Axonal transport of amyloid precursor protein is mediated by direct
RT   binding to the kinesin light chain subunit of kinesin-I.";
RL   Neuron 28:449-459(2000).
RN   [12]
RP   C-TERMINAL PROTEIN-PROTEIN INTERACTIONS, AND MUTAGENESIS OF TYR-728;
RP   THR-743; TYR-757; ASN-759 AND TYR-762.
RX   MEDLINE=21408156; PubMed=11517249;
RA   Matsuda S., Yasukawa T., Homma Y., Ito Y., Niikura T., Hiraki T.,
RA   Hirai S., Ohno S., Kita Y., Kawasumi M., Kouyama K., Yamamoto T.,
RA   Kyriakis J.M., Nishimoto I.;
RT   "C-jun N-terminal kinase (JNK)-interacting protein-1b/islet-brain-1
RT   scaffolds Alzheimer's amyloid precursor protein with JNK.";
RL   J. Neurosci. 21:6597-6607(2001).
RN   [13]
RP   INTERACTION WITH MAPK8IP1, AND PHOSPHORYLATION.
RX   MEDLINE=22028091; PubMed=11912189;
RA   Taru H., Iijima K.-I., Hase M., Kirino Y., Yagi Y., Suzuki T.;
RT   "Interaction of Alzheimer's beta-amyloid precursor family proteins
RT   with scaffold proteins of the JNK signaling cascade.";
RL   J. Biol. Chem. 277:20070-20078(2002).
RN   [14]
RP   INTERACTION OF CTF PEPTIDES WITH NUMB.
RX   MEDLINE=22008109; PubMed=12011466;
RA   Roncarati R., Sestan N., Scheinfeld M.H., Berechid B.E., Lopez P.A.,
RA   Meucci O., McGlade J.C., Rakic P., D'Adamio L.;
RT   "The gamma-secretase-generated intracellular domain of beta-amyloid
RT   precursor protein binds Numb and inhibits Notch signaling.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:7102-7107(2002).
RN   [15]
RP   GAMMA-SECRETASE PROCESSING, AND INTERACTION WITH APBB1.
RX   MEDLINE=21437805; PubMed=11553691;
RA   Cupers P., Orlans I., Craessaerts K., Annaert W., De Strooper B.;
RT   "The amyloid precursor protein (APP)-cytoplasmic fragment generated by
RT   gamma-secretase is rapidly degraded but distributes partially in a
RT   nuclear fraction of neurones in culture.";
RL   J. Neurochem. 78:1168-1178(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions. Can promote transcription activation through binding
CC       to APBB1/Tip60 and inhibit Notch signaling through interaction
CC       with Numb. Couples to apoptosis-inducing pathways such as those
CC       mediated by G(O) and JIP. Inhibits G(O) alpha ATPase activity (By
CC       similarity). Acts as a kinesin I membrane receptor, mediating the
CC       axonal transport of beta-secretase and presenilin 1. May be
CC       involved in copper homeostasis/oxidative stress through copper ion
CC       reduction. Can regulate neurite outgrowth through binding to
CC       components of the extracellular matrix such as heparin and
CC       collagen I and IV (By similarity). The splice isoforms that
CC       contain the BPTI domain possess protease inhibitor activity (By
CC       similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
CC       only weakly transient metals and have little reducing activity due
CC       to substitutions of transient metal chelating residues. Beta-APP42
CC       may activate mononuclear phagocytes in the brain and elicit
CC       inflammatory responses. Promotes both tau aggregation and TPK II-
CC       mediated phosphorylation (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis.
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1, Numb and Dab1. Binding to Dab1 inhibits
CC       its serine phosphorylation. Also interacts with GPCR-like protein
CC       BPP, FPRL1, APPBP1, IB1, KNS2 (via its TPR domains), APPBP2 (via
CC       BaSS) and DDB1 (By similarity). In vitro, it binds MAPT via the
CC       MT-binding domains (By similarity). Associates with microtubules
CC       in the presence of ATP and in a kinesin-dependent manner (By
CC       similarity). Interacts, through a C-terminal domain, with GNAO1
CC       (By similarity). Amyloid beta-42 binds CHRNA7 in hippocampal
CC       neurons (By similarity). Beta-amyloid associates with HADH2 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
```

```
  Query Match             100.0%;  Score 52;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        726 KQYTSIHHG 734
```

```
RESULT 7
A4_PIG
ID   A4_PIG          STANDARD;      PRT;   770 AA.
AC   P79307; Q29023; Q9TUI0;
DT   01-NOV-1997 (Rel. 35, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-
DE   APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);
DE   Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kimura A., Takahashi T.;
```

```
RT   "Amyloid precursor protein 770.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-136 FROM N.A.
RC   TISSUE=Small intestine;
RA   Winteroe A.K., Fredholm M.;
RT   "Evaluation and characterization of a porcine small intestine cDNA
RT   library.";
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 667-723 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=92017079; PubMed=1656157;
RA   Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT   "Conservation of the sequence of the Alzheimer's disease amyloid
RT   peptide in dog, polar bear and five other mammals by cross-species
RT   polymerase chain reaction analysis.";
RL   Brain Res. Mol. Brain Res. 10:299-305(1991).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction (By similarity). In vitro, copper-metallated APP
CC       induces neuronal death directly or is potentiated through Cu(II)-
CC       mediated low-density lipoprotein oxidation (By similarity). Can
CC       regulate neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- PTM: Extracellular binding and reduction of copper, results in a
CC       corresponding oxidation of Cys-144 and Cys-158, and the formation
CC       of a disulfide bond (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ---------------------------------------------------------------------
DR   EMBL; AB032550; BAA84580.1; -.
DR   EMBL; Z84022; CAB06313.1; -.
DR   EMBL; X56127; CAA39592.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Amyloid.
FT   SIGNAL        1     17       BY SIMILARITY.
FT   CHAIN        18    770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN        18    687       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN        18    671       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN       672    770       C99 (BY SIMILARITY).
FT   CHAIN       672    713       BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
FT   CHAIN       672    711       BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
FT   CHAIN       688    770       C83 (BY SIMILARITY).
FT   CHAIN       688    713       P3(42) (BY SIMILARITY).
FT   CHAIN       688    711       P3(40) (BY SIMILARITY).
FT   CHAIN       712    770       GAMMA-CTF(59).
FT   CHAIN       714    770       GAMMA-CTF(57).
FT   CHAIN       721    770       GAMMA-CTF(50) (BY SIMILARITY).
FT   CHAIN       740    770       C31 (DURING APOPTOSIS) (BY SIMILARITY).
FT   DOMAIN       18    699       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    700    723       POTENTIAL.
FT   DOMAIN      724    770       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       96    110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      135    155       COPPER-BINDING (BY SIMILARITY).
FT   DOMAIN      181    188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN      291    341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN      391    423       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      491    522       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      523    540       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN      732    751       INTERACTION WITH G(O)-ALPHA (BY
FT                                SIMILARITY).
FT   DOMAIN      230    260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      274    280       POLY-THR.
FT   SITE        144    144       REQUIRED FOR COPPER(II) REDUCTION
FT                                (BY SIMILARITY).
FT   ACT_SITE    301    302       REACTIVE BOND (BY SIMILARITY).
FT   SITE        671    672       CLEAVAGE (BY BETA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        672    673       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE        687    688       CLEAVAGE (BY ALPHA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        704    704       IMPLICATED IN FREE RADICAL PROPAGATION
FT                                (BY SIMILARITY).
FT   SITE        706    706       INVOLVED IN OXIDATIVE REACTIONS
FT                                (BY SIMILARITY).
FT   SITE        711    712       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                                (BY SIMILARITY).
FT   SITE        713    714       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                                (BY SIMILARITY).
FT   SITE        720    721       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)

  Query Match           100.0%;  Score 52;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        726 KQYTSIHHG 734


RESULT 8
A4_RAT
ID   A4_RAT          STANDARD;      PRT;   770 AA.
AC   P08592;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid
DE   protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains: Soluble
DE   APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99; Beta-
DE   amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40);
DE   C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal
DE   fragment 59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57);
DE   Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50); C31].
GN   APP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=88312583; PubMed=2900758;
RA   Shivers B.D., Hilbich C., Multhaup G., Salbaum J.M., Beyreuther K.,
RA   Seeburg P.H.;
RT   "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
RT   in rat brain suggests a role in cell contact.";
RL   EMBO J. 7:1365-1370(1988).
RN   [2]
RP   SEQUENCE OF 289-364 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=89183625; PubMed=2648331;
RA   Kang J., Mueller-Hill B.;
RT   "The sequence of the two extra exons in rat preA4.";
RL   Nucleic Acids Res. 17:2130-2130(1989).
RN   [3]
RP   SEQUENCE OF 720-730, AND MASS SPECTROMETRY.
```

RX MEDLINE=21443797; PubMed=11483588;
RA Gu Y., Misonou H., Sato T., Dohmae N., Takio K., Ihara Y.;
RT "Distinct intramembrane cleavage of the beta-amyloid precursor protein
RT family resembling gamma-secretase-like cleavage of Notch.";
RL J. Biol. Chem. 276:35235-35238(2001).
RN [4]
RP ALTERNATIVE SPLICING.
RX MEDLINE=96187032; PubMed=8624099;
RA Sandbrink R., Masters C.L., Beyreuther K.;
RT "APP gene family. Alternative splicing generates functionally related
RT isoforms.";
RL Ann. N.Y. Acad. Sci. 777:281-287(1996).
RN [5]
RP TISSUE SPECIFICITY OF APPICAN.
RX MEDLINE=95263526; PubMed=7744833;
RA Shioi J., Pangalos M.N., Ripellino J.A., Vassilacopoulou D.,
RA Mytilineou C., Margolis R.U., Robakis N.K.;
RT "The Alzheimer amyloid precursor proteoglycan (appican) is present in
RT brain and is produced by astrocytes but not by neurons in primary
RT neural cultures.";
RL J. Biol. Chem. 270:11839-11844(1995).
RN [6]
RP TISSUE SPECIFICITY OF ISOFORMS.
RX MEDLINE=97150061; PubMed=8996834;
RA Sandbrink R., Monning U., Masters C.L., Beyreuther K.;
RT "Expression of the APP gene family in brain cells, brain development
RT and aging.";
RL Gerontology 43:119-131(1997).
RN [7]
RP INTERACTION WITH DDB1, AND MUTAGENESIS OF TYR-757; ASN-759 AND
RP TYR-762.
RX MEDLINE=99127916; PubMed=9930726;
RA Watanabe T., Sukegawa J., Tomita S., Iijima K.-I., Oguchi S.,
RA Suzuki T., Nairn A.C., Greengard P.;
RT "A 127-kDa protein (UV-DDB) binds to the cytoplasmic domain of the
RT Alzheimer's amyloid precursor protein.";
RL J. Neurochem. 72:549-556(1999).
RN [8]
RP INTERACTION WITH GNAO1, AND MUTAGENESIS OF 732-HIS-HIS-733.
RX MEDLINE=99162676; PubMed=10024358;
RA Brouillet E., Trembleau A., Galanaud D., Volovitch M., Bouillot C.,
RA Valenza C., Prochiantz A., Allinquant B.;
RT "The amyloid precursor protein interacts with Go heterotrimeric
RT protein within a cell compartment specialized in signal
RT transduction.";
RL J. Neurosci. 19:1717-1727(1999).
RN [9]
RP CHARACTERISTICS OF APPICAN, AND MUTAGENESIS OF SER-656.
RX MEDLINE=95256193; PubMed=7737970;
RA Pangalos M.N., Efthimiopoulos S., Shioi J., Robakis N.K.;
RT "The chondroitin sulfate attachment site of appican is formed by
RT splicing out exon 15 of the amyloid precursor gene.";
RL J. Biol. Chem. 270:10388-10391(1995).
RN [10]
RP BETA-AMYLOID METAL-BINDING.
RX MEDLINE=99316162; PubMed=10386999;
RA Huang X., Atwood C.S., Hartshorn M.A., Multhaup G., Goldstein L.E.,
RA Scarpa R.C., Cuajungco M.P., Gray D.N., Lim J., Moir R.D., Tanzi R.E.,
RA Bush A.I.;
RT "The A beta peptide of Alzheimer's disease directly produces hydrogen
RT peroxide through metal ion reduction.";
RL Biochemistry 38:7609-7616(1999).
RN [11]
RP BETA-AMYLOID ZINC BINDING.
RX MEDLINE=99343552; PubMed=10413512;
RA Liu S.T., Howlett G., Barrow C.J.;
RT "Histidine-13 is a crucial residue in the zinc ion-induced aggregation
RT of the A beta peptide of Alzheimer's disease.";
RL Biochemistry 38:9373-9378(1999).
RN [12]
RP IMPORTANCE OF GLY-704 IN FREE RADICAL PROPAGATION, AND MUTAGENESIS OF
RP GLY-704.
RX MEDLINE=21956095; PubMed=11959460;
RA Kanski J., Varadarajan S., Aksenova M., Butterfield D.A.;
RT "Role of glycine-33 and methionine-35 in Alzheimer's amyloid beta-
RT peptide 1-42-associated oxidative stress and neurotoxicity.";
RL Biochim. Biophys. Acta 1586:190-198(2001).
RN [13]
RP PHOSPHORYLATION.
RX MEDLINE=97239592; PubMed=9085254;
RA Oishi M., Nairn A.C., Czernik A.J., Lim G.S., Isohara T., Gandy S.E.,
RA Greengard P., Suzuki T.;
RT "The cytoplasmic domain of Alzheimer's amyloid precursor protein is
RT phosphorylated at Thr654, Ser655, and Thr668 in adult rat brain and
RT cultured cells.";
RL Mol. Med. 3:111-123(1997).
RN [14]
RP PHOSPHORYLATION ON SER-730.
RX MEDLINE=99262094; PubMed=10329382;
RA Isohara T., Horiuchi A., Watanabe T., Ando K., Czernik A.J., Uno I.,
RA Greengard P., Nairn A.C., Suzuki T.;
RT "Phosphorylation of the cytoplasmic domain of Alzheimer's beta-amyloid
RT precursor protein at Ser655 by a novel protein kinase.";
RL Biochem. Biophys. Res. Commun. 258:300-305(1999).
RN [15]
RP PHOSPHORYLATION, INDUCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
RP THR-743.
RX MEDLINE=99274744; PubMed=10341243;
RA Ando K., Oishi M., Takeda S., Iijima K.-I., Isohara T., Nairn A.C.,
RA Kirino Y., Greengard P., Suzuki T.;
RT "Role of phosphorylation of Alzheimer's amyloid precursor protein
RT during neuronal differentiation.";
RL J. Neurosci. 19:4421-4427(1999).
RN [16]
RP PHOSPHORYLATION ON THR-743.
RX MEDLINE=20396183; PubMed=10936190;
RA Iijima K.-I., Ando K., Takeda S., Satoh Y., Seki T., Itohara S.,
RA Greengard P., Kirino Y., Nairn A.C., Suzuki T.;
RT "Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor
RT protein by cyclin-dependent kinase 5.";
RL J. Neurochem. 75:1085-1091(2000).
RN [17]
RP CARBOHYDRATE STRUCTURE OF APPICAN.
RX MEDLINE=21463085; PubMed=11479316;

```
RA   Tsuchida K., Shioi J., Yamada S., Boghosian G., Wu A., Cai H.,
RA   Sugahara K., Robakis N.K.;
RT   "Appican, the proteoglycan form of the amyloid precursor protein,
RT   contains chondroitin sulfate E in the repeating disaccharide region
RT   and 4-O-sulfated galactose in the linkage region.";
RL   J. Biol. Chem. 276:37155-37160(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP. Inhibits
CC       G(O) alpha ATPase activity. Acts as a kinesin I membrane receptor,
CC       mediating the axonal transport of beta-secretase and presenilin 1
CC       (By similarity). May be involved in copper homeostasis/oxidative
CC       stress through copper ion reduction. Can regulate neurite
CC       outgrowth through binding to components of the extracellular
CC       matrix such as heparin and collagen I and IV (By similarity). The
CC       splice isoforms that contain the BPTI domain possess protease
CC       inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
CC       only weakly transient metals and have little reducing activity due
CC       to substitutions of transient metal chelating residues. Beta-APP42
CC       may activate mononuclear phagocytes in the brain and elicit
CC       inflammatory responses. Promotes both tau aggregation and TPK II-
CC       mediated phosphorylation (By similarity).
CC   -!- FUNCTION: Appicans elicit adhesion of neural cells to the
CC       extracellular matrix and may regulate neurite outgrowth in the
CC       brain.
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains), APPBP2 (via BaSS) (By similarity) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity). Interacts,
CC       through a C-terminal domain, with GNAO1. Amyloid beta-42 binds
CC       CHRNA7 in hippocampal neurons (By similarity). Beta-amyloid
CC       associates with HADH2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the

  Query Match             100.0%;  Score 52;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9

                 |||||||||
Db        726 KQYTSIHHG 734


RESULT 9
A4_TETFL
ID   A4_TETFL       STANDARD;      PRT;   780 AA.
AC   O73683;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Alzheimer's disease amyloid A4 protein homolog precursor [Contains:
DE   Beta-amyloid protein (Beta-APP) (A-beta)].
GN   APP.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98252138; PubMed=9599080;
RA   Villard L., Tassone F., Crnogorac-Jurcevic T., Clancy K., Gardiner K.;
RT   "Analysis of pufferfish homologues of the AT-rich human APP gene.";
RL   Gene 210:17-24(1998).
CC   -!- FUNCTION: Functional neuronal receptor which couples to
CC       intracellular signaling pathway through the GTP-binding protein
CC       G(O) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF018165; AAC41275.1; -.
DR   HSSP; P05067; 1HZ3.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
```

```
DR   PROSITE; PS00280; BPTI_KUNITZ_1; FALSE_NEG.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
KW   Serine protease inhibitor.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    780       ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
FT                                 HOMOLOG.
FT   CHAIN       682    724       BETA-AMYLOID PROTEIN (POTENTIAL).
FT   DOMAIN       19    711       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    712    732       POTENTIAL.
FT   DOMAIN      733    780       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      323    382       BPTI/KUNITZ INHIBITOR.
FT   SITE        769    772       CLATHRIN-BINDING (BY SIMILARITY).
FT   DISULFID    327    378       BY SIMILARITY.
FT   DISULFID    336    361       BY SIMILARITY.
FT   CARBOHYD    560    560       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   780 AA;  88238 MW;  60071BE94520191D CRC64;

  Query Match             100.0%;  Score 52;  DB 1;  Length 780;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        736 KQYTSIHHG 744

Search completed: October  4, 2004, 18:57:05
Job time : 10.4348 secs
```

```
                                GenCore version 5.1.6
                          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October  4, 2004, 18:37:37 ; Search time 42.6522 Seconds
                                              (without alignments)
                                              66.577 Million cell updates/sec

Title:           US-10-048-209-2
Perfect score:   52
Sequence:        1 KQYTSIHHG 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:      1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        SPTREMBL_25:*
                 1:  sp_archea:*
                 2:  sp_bacteria:*
                 3:  sp_fungi:*
                 4:  sp_human:*
                 5:  sp_invertebrate:*
                 6:  sp_mammal:*
                 7:  sp_mhc:*
                 8:  sp_organelle:*
                 9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_virus:*
                 13:  sp_vertebrate:*
                 14:  sp_unclassified:*
                 15:  sp_rvirus:*
                 16:  sp_bacteriap:*
                 17:  sp_archeap:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                    SUMMARIES

                                    %
Result                              Query
   No.    Score                     Match Length DB  ID                      Description
-------------------------------------------------------------------------------------------
```

1	52	100.0	49	6	O97917	O97917 bos taurus
2	52	100.0	113	13	Q8JH58	Q8jh58 chelydra se
3	52	100.0	218	11	Q8BPV5	Q8bpv5 mus musculu
4	52	100.0	239	13	Q8UUI7	Q8uui7 brachydanio
5	52	100.0	357	13	Q8UUI8	Q8uui8 brachydanio
6	52	100.0	384	11	Q8BPC7	Q8bpc7 mus musculu
7	52	100.0	472	13	Q8UUS0	Q8uus0 brachydanio
8	52	100.0	534	13	O93296	O93296 gallus gall
9	52	100.0	569	13	Q9PVL1	Q9pvl1 gallus gall
10	52	100.0	612	13	Q9I9E7	Q9i9e7 brachydanio
11	52	100.0	678	13	Q7ZZT1	Q7zzt1 brachydanio
12	52	100.0	694	13	Q8UUR9	Q8uur9 brachydanio
13	52	100.0	695	13	Q9DGJ8	Q9dgj8 gallus gall
14	52	100.0	699	13	O57394	O57394 narke japon
15	52	100.0	738	13	Q90W28	Q90w28 brachydanio
16	52	100.0	751	13	Q9DGJ7	Q9dgj7 gallus gall
17	49	94.2	693	13	Q98SG0	Q98sg0 xenopus lae
18	49	94.2	695	13	Q98SF9	Q98sf9 xenopus lae
19	49	94.2	695	13	Q7ZXQ0	Q7zxq0 xenopus lae
20	49	94.2	747	13	Q91963	Q91963 xenopus. ap
21	45	86.5	82	4	Q16020	Q16020 homo sapien
22	45	86.5	82	4	Q16014	Q16014 homo sapien
23	45	86.5	82	4	Q16019	Q16019 homo sapien
24	37	71.2	133	5	O17443	O17443 brevicoryne
25	37	71.2	164	16	Q81P90	Q81p90 bacillus an
26	37	71.2	228	16	Q8A0P0	Q8a0p0 bacteroides
27	37	71.2	442	16	Q92AT6	Q92at6 listeria in
28	37	71.2	2584	5	Q9TW88	Q9tw88 caenorhabdi
29	37	71.2	2606	5	Q21920	Q21920 caenorhabdi
30	36	69.2	71	16	Q83DX8	Q83dx8 coxiella bu
31	36	69.2	116	11	Q8CDE5	Q8cde5 mus musculu
32	36	69.2	159	17	Q8TTR5	Q8ttr5 methanosarc
33	36	69.2	159	17	Q8TTR3	Q8ttr3 methanosarc
34	36	69.2	275	16	Q9RXF2	Q9rxf2 deinococcus
35	36	69.2	333	13	Q90YZ9	Q90yz9 sphyraena i
36	36	69.2	412	5	Q21806	Q21806 caenorhabdi
37	36	69.2	443	2	Q9ZA32	Q9za32 streptomyce
38	36	69.2	550	8	Q7YAJ6	Q7yaj6 chara vulga
39	36	69.2	729	16	Q8DKG0	Q8dkg0 synechococc
40	36	69.2	5058	4	Q86UQ4	Q86uq4 homo sapien
41	35	67.3	130	2	Q9XBK5	Q9xbk5 bacillus ce
42	35	67.3	150	16	Q97MJ3	Q97mj3 clostridium
43	35	67.3	210	11	Q7TPI2	Q7tpi2 peromyscus
44	35	67.3	251	13	Q8AWR6	Q8awr6 brachydanio
45	35	67.3	277	16	Q81VK1	Q81vk1 bacillus an

ALIGNMENTS

RESULT 1
O97917
ID O97917 PRELIMINARY; PRT; 49 AA.
AC O97917;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Amyloid protein (Fragment).
GN APP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20063685; PubMed=10594237;
RA Konfortov B.A., Licence V.E., Miller J.R.;
RT "Resequencing of DNA from a diverse panel of cattle reveals a high
RT level of polymorphism in both intron and exon.";
RL Mamm. Genome 10:1142-1145(1999).
DR EMBL; AJ133033; CAB38017.1; -.
DR HSSP; P05067; 1BA4.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.
FT NON_TER 1 1
FT NON_TER 49 49
SQ SEQUENCE 49 AA; 5183 MW; 6287463F0559BDED CRC64;

  Query Match 100.0%; Score 52; DB 6; Length 49;
  Best Local Similarity 100.0%; Pred. No. 0.0092;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQYTSIHHG 9
    |||||||||
Db 38 KQYTSIHHG 46

RESULT 2
Q8JH58
ID Q8JH58 PRELIMINARY; PRT; 113 AA.
AC Q8JH58;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Amyloid beta protein (Fragment).
OS Chelydra serpentina serpentina (common snapping turtle).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Cryptodira; Testudinoidea; Chelydridae; Chelydra.
OX NCBI_TaxID=134619;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21876906; PubMed=11882478;
RA Trudeau V.L., Chiu S., Kennedy S.W., Brooks R.J.;
RT "Octylphenol (OP) alters the expression of members of the amyloid
RT protein family in the hypothalamus of the snapping turtle, Chelydra
RT serpentina serpentina.";
RL Environ. Health Perspect. 110:269-275(2002).
DR EMBL; AF541917; AAN04908.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR008155; A4_APP.
DR InterPro; IPR001255; Beta-APP.
DR Pfam; PF03494; Beta-APP; 1.

```
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   113 AA;  12750 MW;  72515C930496E053 CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 113;
  Best Local Similarity   100.0%;  Pred. No. 0.022;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db         69 KQYTSIHHG 77


RESULT 3
Q8BPV5
ID   Q8BPV5      PRELIMINARY;      PRT;   218 AA.
AC   Q8BPV5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid beta (Fragment).
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK052448; BAC34997.1; -.
DR   MGD; MGI:88059; App.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   218 AA;  24118 MW;  95B55AFDAE1D0EF5 CRC64;

  Query Match             100.0%;  Score 52;  DB 11;  Length 218;
  Best Local Similarity   100.0%;  Pred. No. 0.044;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        174 KQYTSIHHG 182


RESULT 4
Q8UUI7
ID   Q8UUI7      PRELIMINARY;      PRT;   239 AA.
AC   Q8UUI7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative membrane protein (Fragment).
GN   APPB.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315638; CAC85735.1; -.
DR   ZFIN; ZDB-GENE-020220-1; appb.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   239 AA;  27048 MW;  8A69F746F821BAE2 CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 0.049;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        195 KQYTSIHHG 203


RESULT 5
Q8UUI8
ID   Q8UUI8      PRELIMINARY;      PRT;   357 AA.
AC   Q8UUI8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative mebrane protein (Fragment).
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Embryo;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315637; CAC85734.1; -.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   357 AA;  40962 MW;  07D99EEF6C55B2D8 CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 357;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        313 KQYTSIHHG 321


RESULT 6
Q8BPC7
ID   Q8BPC7      PRELIMINARY;      PRT;   384 AA.
AC   Q8BPC7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid beta (Fragment).
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK076506; BAC36369.1; -.
DR   MGD; MGI:88059; App.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   384 AA;  43990 MW;  A81B1AD8AE683173 CRC64;

  Query Match             100.0%;  Score 52;  DB 11;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 0.08;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        340 KQYTSIHHG 348


RESULT 7
Q8UUS0
ID   Q8UUS0      PRELIMINARY;      PRT;   472 AA.
AC   Q8UUS0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative membrane protein (Fragment).
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315636; CAC85733.1; -.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   472 AA;  53787 MW;  24F7128BE3356550 CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 0.1;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        428 KQYTSIHHG 436


RESULT 8
O93296
ID   O93296      PRELIMINARY;      PRT;   534 AA.
AC   O93296;
```

```
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98337885; PubMed=9671674;
RA   Barnes N.Y., Li L., Yoshikawa K., Schwartz L.M., Oppenheim R.W.,
RA   Milligan C.E.;
RT   "Increased production of amyloid precursor protein provides a
RT   substrate for caspase-3 in dying motoneurons.";
RL   J. Neurosci. 18:5869-5880(1998).
DR   EMBL; AF042098; AAC25052.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   534 AA;  60597 MW;  FB53ECC2E66D4C92 CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 0.11;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        490 KQYTSIHHG 498


RESULT 9
Q9PVL1
ID   Q9PVL1      PRELIMINARY;      PRT;   569 AA.
AC   Q9PVL1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
GN   APP.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Coulson E.J., Paliga K., Beyreuther K., Masters C.L.;
RT   "What the evolution of the amyloid protein precursor supergene family
RT   tells us about its function.";
RL   Neurochem. Int. 0:0-0(2000).
DR   EMBL; AF030341; AAF12698.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   569 AA;  64753 MW;  0AB8BB851863A19D CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 569;
  Best Local Similarity   100.0%;  Pred. No. 0.12;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        526 KQYTSIHHG 534


RESULT 10
Q9I9E7
ID   Q9I9E7      PRELIMINARY;      PRT;   612 AA.
AC   Q9I9E7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Slavov D.B., Gardiner K.;
RT   "An App cDNA from Zebrafish (Danio rerio).";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF257742; AAF71748.1; -.
DR   HSSP; P05067; 1HZ3.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
```

```
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   612 AA;  69710 MW;  59A9ACBDF9C59EFF CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 612;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        568 KQYTSIHHG 576


RESULT 11
Q7ZZT1
ID   Q7ZZT1      PRELIMINARY;      PRT;   678 AA.
AC   Q7ZZT1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid protein a variant 2.
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Groth C., Lardelli M.;
RT   "Investigation of zebrafish appa expression during embryogenesis.";
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY271746; AAP22958.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   678 AA;  76755 MW;  94163778444FD0BC CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 678;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        634 KQYTSIHHG 642


RESULT 12
Q8UUR9
ID   Q8UUR9      PRELIMINARY;      PRT;   694 AA.
AC   Q8UUR9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative membrane protein.
GN   APPB.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315639; CAC85736.1; -.
DR   ZFIN; ZDB-GENE-020220-1; appb.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   694 AA;  79228 MW;  2B03382D411162DC CRC64;

  Query Match             100.0%;  Score 52;  DB 13;  Length 694;
  Best Local Similarity   100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        650 KQYTSIHHG 658


RESULT 13
Q9DGJ8
ID   Q9DGJ8      PRELIMINARY;      PRT;   695 AA.
AC   Q9DGJ8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Beta-amyloid precursor protein 695 isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Sarasa M., Rodolosse A., Sorribas V.;
RT   "Cloning of full-length chicken beta-amyloid precursor protein
RT   isoforms.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289218; AAG00593.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   695 AA;  78565 MW;  F201ED02AEC86D95 CRC64;

  Query Match           100.0%;  Score 52;  DB 13;  Length 695;
  Best Local Similarity  100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        651 KQYTSIHHG 659


RESULT 14
O57394
ID   O57394      PRELIMINARY;      PRT;   699 AA.
AC   O57394;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   EL amyloid precursor protein 699.
GN   EL APP699.
OS   Narke japonica (Electric ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Torpediniformes; Narcinoidei; Narkidae; Narke.
OX   NCBI_TaxID=62965;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Electric lobe;
RX   MEDLINE=98129705; PubMed=9461486;
RA   Iijima K., Lee D.-S., Okutsu J., Tomita S., Hirashima N., Kirino Y.,
RA   Suzuki T.;
RT   "cDNA isolation of Alzheimer's amyloid precursor protein from
RT   cholinergic nerve terminals of the electric organ of the electric
RT   ray.";
RL   Biochem. J. 330:29-33(1998).
DR   EMBL; AB005544; BAA24230.1; -.
DR   HSSP; P05067; 1HZ3.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   699 AA;  78879 MW;  952915C309D50E5C CRC64;

  Query Match           100.0%;  Score 52;  DB 13;  Length 699;
  Best Local Similarity  100.0%;  Pred. No. 0.15;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQYTSIHHG 9
              |||||||||
Db        655 KQYTSIHHG 663


RESULT 15
Q90W28
ID   Q90W28      PRELIMINARY;      PRT;   738 AA.
AC   Q90W28;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid precursor protein.
GN   APPA OR APP.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Groth C., Lardelli M.;
RT   "Expression analysis of zebrafish app.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF389401; AAK64495.1; -.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Protease inhibitor; Serine protease inhibitor.
```

SQ SEQUENCE 738 AA; 83577 MW; AF480F6D308FD298 CRC64;

  Query Match 100.0%; Score 52; DB 13; Length 738;
  Best Local Similarity 100.0%; Pred. No. 0.16;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQYTSIHHG 9
              |||||||||
Db        694 KQYTSIHHG 702
```

Search completed: October 4, 2004, 18:56:21
Job time : 46.6522 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:37:13 ; Search time 58.0435 Seconds
(without alignments)
48.679 Million cell updates/sec

Title: US-10-048-209-3
Perfect score: 57
Sequence: 1 KKQYTSIHHG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	100.0	12	4	AAB67793	Aab67793 Cytoplasm
2	57	100.0	13	4	AAB67794	Aab67794 Cytoplasm
3	57	100.0	15	4	AAB67798	Aab67798 Cytoplasm
4	57	100.0	19	7	ADC49255	Adc49255 Human inh
5	57	100.0	32	6	AAO19883	Aao19883 Human amy
6	57	100.0	33	7	ADC49254	Adc49254 Human inh
7	57	100.0	41	4	AAM16658	Aam16658 Peptide #
8	57	100.0	41	4	ABB35642	Abb35642 Peptide #
9	57	100.0	41	4	AAM29142	Aam29142 Peptide #
10	57	100.0	41	4	ABB30475	Abb30475 Peptide #
11	57	100.0	41	4	ABB21071	Abb21071 Protein #
12	57	100.0	41	4	AAM56458	Aam56458 Human bra
13	57	100.0	41	4	AAM04374	Aam04374 Peptide #
14	57	100.0	41	5	ABG38416	Abg38416 Human pep
15	57	100.0	44	2	AAW53985	Aaw53985 Human ALZ
16	57	100.0	47	2	AAR58917	Aar58917 Cytoplasm
17	57	100.0	47	2	AAW26395	Aaw26395 Amyloid p
18	57	100.0	47	2	AAW26402	Aaw26402 Amyloid p
19	57	100.0	47	2	AAW26400	Aaw26400 Amyloid p
20	57	100.0	47	2	AAW26403	Aaw26403 Amyloid p
21	57	100.0	47	2	AAW26399	Aaw26399 Amyloid p
22	57	100.0	47	2	AAW26401	Aaw26401 Amyloid p
23	57	100.0	47	2	AAW26520	Aaw26520 Amyloid p
24	57	100.0	47	2	AAW26518	Aaw26518 Amyloid p
25	57	100.0	47	2	AAW26521	Aaw26521 Amyloid p
26	57	100.0	47	2	AAW26519	Aaw26519 Amyloid p
27	57	100.0	47	2	AAW26513	Aaw26513 Amyloid p
28	57	100.0	47	2	AAW26517	Aaw26517 Amyloid p
29	57	100.0	47	2	AAW42984	Aaw42984 APP isofo
30	57	100.0	47	2	AAW42986	Aaw42986 APP isofo
31	57	100.0	47	2	AAW42987	Aaw42987 APP isofo
32	57	100.0	47	2	AAW42983	Aaw42983 APP isofo
33	57	100.0	47	2	AAW42985	Aaw42985 APP isofo
34	57	100.0	47	2	AAW44755	Aaw44755 APP-REP 7
35	57	100.0	47	2	AAW44753	Aaw44753 APP-REP 7
36	57	100.0	47	2	AAW44749	Aaw44749 APP-REP 7
37	57	100.0	47	2	AAW44756	Aaw44756 APP-REP 7
38	57	100.0	47	2	AAW44754	Aaw44754 APP-REP 7
39	57	100.0	47	2	AAW44757	Aaw44757 APP-REP 7
40	57	100.0	47	2	AAW44274	Aaw44274 Amyloid p
41	57	100.0	47	4	AAB67791	Aab67791 Cytoplasm
42	57	100.0	47	7	ADC49250	Adc49250 Human inh
43	57	100.0	47	7	ADC49249	Adc49249 Inhibitor
44	57	100.0	47	7	ADC49251	Adc49251 Human inh
45	57	100.0	49	2	AAR35087	Aar35087 Human amy

ALIGNMENTS

RESULT 1
AAB67793
ID AAB67793 standard; peptide; 12 AA.
XX
AC AAB67793;
XX
DT 11-JUN-2001 (first entry)
XX
DE Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1

```
FT                   /note= "this residue represents an internalisation
FT                   peptide such as the sequence given in AAB67795"
FT   Misc-difference 12
FT                   /note= "this residue represents V, VV, VVE, VVEV, VVEVD"
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Claim 1; Page 13; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 12 AA;

  Query Match             100.0%;  Score 57;  DB 4;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.00082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db          2 KKQYTSIHHG 11


RESULT 2
AAB67794
ID   AAB67794 standard; peptide; 13 AA.
XX
AC   AAB67794;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW   Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
```

```
FT   Misc-difference 1
FT                   /note= "this residue represents an internalisation
FT                   peptide such as the sequence given in AAB67795"
FT   Misc-difference 13
FT                   /note= "this residue represents V, VV, VVE, VVEV, VVEVD"
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Claim 1; Page 13; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 57;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00089;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db          3 KKQYTSIHHG 12


RESULT 3
AAB67798
ID   AAB67798 standard; peptide; 15 AA.
XX
AC   AAB67798;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW   Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
```

PN WO200109170-A1.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-FR002174.
XX
PR 30-JUL-1999; 99FR-00009929.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Allinquant B, Prochiantz A;
XX
DR WPI; 2001-257398/26.
XX
PT Peptides derived from the cytoplasmic domain of the amyloid protein
PT precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS Disclosure; Page 2; 28pp; French.
XX
CC The present sequence represents a peptide derived from the cytoplasmic
CC domain of the human amyloid protein precursor (APP). APP peptides derived
CC from the cytoplasmic domain, and containing the membrane domain
CC juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC screening products capable of inhibiting apoptosis. The peptides are
CC useful in the treatment of cancer and Alzheimer's disease
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 57; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 2 KKQYTSIHHG 11

RESULT 4
ADC49255
ID ADC49255 standard; protein; 19 AA.
XX
AC ADC49255;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human inhibitor of metabolic degradation of APP variant #4.
XX
KW human; inhibitory factor; metabolic degradation of APP;
KW amyloid precursor protein; Alzheimer's disease; mutant; mutein.
XX
OS Homo sapiens.
XX
PN JP2002360252-A.
XX
PD 17-DEC-2002.
XX
PF 27-APR-2001; 2001JP-00133178.
XX
PR 27-APR-2001; 2001JP-00133178.
XX
PA (SUZU/) SUZUKI T.
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2003-516151/49.
XX
PT An inhibitory factor of metabolic degradation of amyloid precursor
PT protein (APP) which inhibits formation of beta-amyloid, useful in the
PT treatment of Alzheimer's disease.
XX
PS Disclosure; SEQ ID NO 15; 33pp; Japanese.
XX
CC The invention relates to an inhibitory factor of metabolic degradation of
CC APP. The factor is useful in the treatment of Alzheimer's disease. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 19 AA;

Query Match 100.0%; Score 57; DB 7; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 2 KKQYTSIHHG 11

RESULT 5
AAO19883
ID AAO19883 standard; peptide; 32 AA.
XX
AC AAO19883;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human amyloid precursor protein APP immunogenic peptide #3.
XX
KW Human; APP; amyloid precursor protein; immunogen; Alzheimer's disease;
KW high-throughput screening; neuroprotective; nootropic; antiparkinsonian.
XX
OS Homo sapiens.
XX
PN WO2003001881-A2.
XX
PD 09-JAN-2003.
XX
PF 26-JUN-2002; 2002WO-US020267.
XX
PR 26-JUN-2001; 2001US-0300959P.
XX
PA (NYME-) NEW YORK STATE OFFICE MENTAL HEALTH.
XX
PI Mathews PM, Nixon RA, Schmidt SD, Jiang Y;
XX
DR WPI; 2003-210182/20.

XX
PT Identifying compounds that modulates the generation of metabolites
PT associated with a disease or disorder, for treating e.g. Alzheimer's
PT disease by determining levels of a cellular component protein, or its
PT conformation state.
XX
PS Example 1; Page 29; 69pp; English.
XX
CC The present invention relates to a method of identifying compounds that
CC modulate the generation of one or more metabolites associated with a
CC disease or disorder comprising determining levels of a cellular component
CC protein or a conformation state of a cellular precursor protein. In
CC particular, the method can be used to determine levels of amyloid
CC precursor protein (APP), which is associated with Alzheimer's disease. It
CC is also useful for identifying compounds as drugs for treating diseases
CC or disorders associated with metabolic and/or proteolytic pathways, e.g.
CC Alzheimer's disease, Parkinson's disease, Huntington's disease, lysosomal
CC storage disorders, prion diseases, the tau-based neurodegenerative
CC disorders, and other non-AD amyloidoses. The present sequence is an
CC immunogenic portion of human APP
XX
SQ Sequence 32 AA;

Query Match 100.0%; Score 57; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 6 KKQYTSIHHG 15

RESULT 6
ADC49254
ID ADC49254 standard; protein; 33 AA.
XX
AC ADC49254;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human inhibitor of metabolic degradation of APP variant #3.
XX
KW human; inhibitory factor; metabolic degradation of APP;
KW amyloid precursor protein; Alzheimer's disease; mutant; mutein.
XX
OS Homo sapiens.
XX
PN JP2002360252-A.
XX
PD 17-DEC-2002.
XX
PF 27-APR-2001; 2001JP-00133178.
XX
PR 27-APR-2001; 2001JP-00133178.
XX
PA (SUZU/) SUZUKI T.
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2003-516151/49.
XX
PT An inhibitory factor of metabolic degradation of amyloid precursor
PT protein (APP) which inhibits formation of beta-amyloid, useful in the
PT treatment of Alzheimer's disease.
XX
PS Disclosure; SEQ ID NO 14; 33pp; Japanese.
XX
CC The invention relates to an inhibitory factor of metabolic degradation of
CC APP. The factor is useful in the treatment of Alzheimer's disease. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 57; DB 7; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 2 KKQYTSIHHG 11

RESULT 7
AAM16658
ID AAM16658 standard; protein; 41 AA.
XX
AC AAM16658;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #3092 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 27; SEQ ID NO 21484; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 57; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 29 KKQYTSIHHG 38

RESULT 8
ABB35642
ID ABB35642 standard; peptide; 41 AA.
XX
AC ABB35642;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #3148 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 28277; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 57; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 29 KKQYTSIHHG 38

RESULT 9
AAM29142
ID AAM29142 standard; protein; 41 AA.
XX
AC AAM29142;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #3179 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 29411; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 57; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 29 KKQYTSIHHG 38

RESULT 10
ABB30475
ID ABB30475 standard; peptide; 41 AA.
XX
AC ABB30475;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #3126 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 13443; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 57; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
||||||||||
Db 29 KKQYTSIHHG 38

RESULT 11
ABB21071
ID ABB21071 standard; protein; 41 AA.
XX
AC ABB21071;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #3070 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 22841; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 57; DB 4; Length 41;
  Best Local Similarity 100.0%; Pred. No. 0.0031;
  Matches   10; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKQYTSIHHG 10
              ||||||||||
Db   29 KKQYTSIHHG 38

RESULT 12
AAM56458
ID AAM56458 standard; protein; 41 AA.
XX
AC AAM56458;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28563.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 28563; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 57; DB 4; Length 41;
  Best Local Similarity 100.0%; Pred. No. 0.0031;
  Matches   10; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKQYTSIHHG 10
              ||||||||||
Db   29 KKQYTSIHHG 38

RESULT 13

AAM04374
ID AAM04374 standard; protein; 41 AA.
XX
AC AAM04374;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #3056 encoded by probe for measuring breast gene expression.
XX
KW Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 27; SEQ ID NO 13114; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

 Query Match   100.0%; Score 57; DB 4; Length 41;
 Best Local Similarity   100.0%; Pred. No. 0.0031;
  Matches   10; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKQYTSIHHG 10
              ||||||||||
Db   29 KKQYTSIHHG 38

RESULT 14
ABG38416
ID ABG38416 standard; peptide; 41 AA.
XX
AC ABG38416;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 28081.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
PS Claim 27; SEQ ID NO 28081; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of

```
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a peptide/protein encoded by a single exon probe of
CC   the invention. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 41 AA;

  Query Match           100.0%;  Score 57;  DB 5;  Length 41;
  Best Local Similarity  100.0%;  Pred. No. 0.0031;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db         29 KKQYTSIHHG 38


RESULT 15
AAW53985
ID   AAW53985 standard; protein; 44 AA.
XX
AC   AAW53985;
XX
DT   18-AUG-1998  (first entry)
XX
DE   Human ALZASp2.
XX
KW   Dsas; DSASp; alzas; Down's syndrome; diagnosis; therapy; human;
KW   Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO9807850-A2.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-EP004599.
XX
PR   22-AUG-1996;   96CA-02183901.
XX
PA   (BERG/) BERGMANN J E.
PA   (PRED/) PREDDIE E R.
XX
PI   Bergmann JE,  Preddie ER;
XX
DR   WPI; 1998-169155/15.
DR   N-PSDB; AAV23755.
XX
PT   Nucleic acid molecules dsas, and alzas - used for detecting and treating
PT   Down's syndrome and Alzheimer's disease.
XX
PS   Claim 13; Fig 1M; 96pp; English.
XX
CC   This sequence is the ALZASp2 encoded by the nucleic acid alzas. The dsas
CC   and alzas DNA sequences are the nucleic acids of the invention. Reagents
CC   specifically for DSASp can be used for the diagnosis of Down's syndrome
CC   in humans and especially in pregnant women. Molecules that inhibit the
CC   activity of the promoters (PDS1, PDS2, PDS3, and PDS4) for dsas can be
CC   used for treating Down's syndrome. The reagent capable of detecting alzas
CC   can be used for detecting Alzheimer's disease, especially in the pre-
CC   symptomatic stage. Substances that inhibit the promoters for alzas can be
CC   used in treating Alzheimer's disease
XX
SQ   Sequence 44 AA;

  Query Match           100.0%;  Score 57;  DB 2;  Length 44;
  Best Local Similarity  100.0%;  Pred. No. 0.0033;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db         20 KKQYTSIHHG 29


Search completed: October  4, 2004, 18:52:36
Job time : 58.0435 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:45:55 ; Search time 14.7826 Seconds
(without alignments)
65.071 Million cell updates/sec

Title: US-10-048-209-3
Perfect score: 57
Sequence: 1 KKQYTSIHHG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	100.0	82	2	PQ0438	Alzheimer's diseas
2	57	100.0	695	1	A49795	Alzheimer's diseas
3	57	100.0	695	2	A27485	Alzheimer's diseas
4	57	100.0	695	2	S00550	Alzheimer's diseas
5	57	100.0	770	1	QRHUA4	Alzheimer's diseas
6	54	94.7	747	2	JH0773	Alzheimer's diseas
7	39	68.4	452	2	S58994	NADH2 dehydrogenas
8	38	66.7	406	2	S76307	hypothetical prote
9	38	66.7	488	2	S64140	hypothetical prote
10	37	64.9	142	2	T10871	y4kQ protein - Rhi
11	37	64.9	442	2	AH1661	ATP-dependent RNA
12	37	64.9	2584	2	T24158	hypothetical prote
13	37	64.9	2606	2	T24157	hypothetical prote
14	36	63.2	275	2	C75529	competence protein
15	36	63.2	412	2	T24023	hypothetical prote
16	36	63.2	443	2	T46532	probable dTDP-4-ke
17	36	63.2	1268	2	B88209	protein K02A2.6 [i
18	36	63.2	1313	2	T29193	hypothetical prote
19	35	61.4	150	2	F96924	flavodoxin [import
20	35	61.4	191	2	A35981	sperm membrane pro
21	35	61.4	239	2	T01463	hypothetical prote
22	35	61.4	264	2	T15289	hypothetical prote
23	35	61.4	410	2	AF1660	aminopeptidases ho
24	35	61.4	410	2	AG1288	aminopeptidases ho
25	35	61.4	443	2	B39794	transcription fact
26	35	61.4	444	1	A39794	transcription fact
27	35	61.4	452	2	T04781	hypothetical prote
28	35	61.4	511	2	JC1404	CDEI-box DNA-bindi
29	35	61.4	522	2	A83804	involved in spore
30	35	61.4	739	2	AG1667	phosphoribosylform
31	35	61.4	739	2	AI1295	phosphoribosylform
32	35	61.4	751	2	A49974	beta-amyloid precu
33	35	61.4	763	2	A49321	amyloid beta (A4)
34	35	61.4	765	2	S42880	amyloid precursor-
35	35	61.4	848	2	E85087	hypothetical prote
36	34	59.6	99	2	B82524	hypothetical prote
37	34	59.6	100	2	T41338	very hypothetical
38	34	59.6	119	2	AG3136	hypothetical prote
39	34	59.6	254	2	T29556	hypothetical prote
40	34	59.6	332	2	T25779	hypothetical prote
41	34	59.6	344	1	TWXL3	transcription fact
42	34	59.6	348	2	S23900	nitrogen regulatio
43	34	59.6	393	2	C96592	hypothetical prote
44	34	59.6	409	2	C96956	permease [imported
45	34	59.6	637	2	H70535	hypothetical prote

ALIGNMENTS

RESULT 1
PQ0438
Alzheimer's disease amyloid A4 protein precursor - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C;Accession: PQ0438; C60045
R;Davidson, J.S.; West, R.L.; Kotikalapudi, P.; Maroun, L.E.
Biochem. Biophys. Res. Commun. 188, 905-911, 1992
A;Title: Sequence and methylation in the beta/A4 region of the rabbit amyloid precursor protein gene.
A;Reference number: PQ0438; MUID:93075180; PMID:1445331
A;Accession: PQ0438
A;Molecule type: DNA
A;Residues: 1-82 <DAV>
A;Cross-references: GB:M83558; GB:M83657
R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157
A;Accession: C60045
A;Molecule type: mRNA
A;Residues: 12-68 <JOH>
A;Cross-references: EMBL:X56129
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; Alzheimer's disease; amyloid; Down's syndrome

  Query Match 100.0%; Score 57; DB 2; Length 82;
  Best Local Similarity 100.0%; Pred. No. 0.00089;
  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
    ||||||||||
Db 70 KKQYTSIHHG 79

RESULT 2
A49795
Alzheimer's disease amyloid beta protein precursor - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A49795
R;Podlisny, M.B.; Tolan, D.R.; Selkoe, D.J.
Am. J. Pathol. 138, 1423-1435, 1991
A;Title: Homology of the amyloid beta protein precursor in monkey and human supports a primate model for beta amyloidosis in Alzheimer's disease.
A;Reference number: A49795; MUID:91273117; PMID:1905108
A;Accession: A49795
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-695 <POD>
A;Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing

  Query Match 100.0%; Score 57; DB 1; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.0089;
  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
    ||||||||||
Db 650 KKQYTSIHHG 659

RESULT 3
A27485
Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse
N;Alternate names: proteinase nexin II
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999
C;Accession: A27485; S19727; I49485
R;Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sakaki, Y.
Biochem. Biophys. Res. Commun. 149, 665-671, 1987
A;Title: Complementary DNA for the mouse homolog of the human amyloid beta protein precursor.
A;Reference number: A27485; MUID:88106489; PMID:3322280
A;Accession: A27485
A;Molecule type: mRNA
A;Residues: 1-695 <YAM>
A;Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085
A;Experimental source: brain
R;de Strooper, B.; van Leuven, F.; van den Berghe, H.
Biochim. Biophys. Acta 1129, 141-143, 1991
A;Title: The amyloid beta protein precursor or proteinase nexin II from mouse is closer related to its human homolog than previously reported.
A;Reference number: S19727; MUID:92096458; PMID:1756177
A;Accession: S19727
A;Molecule type: mRNA
A;Residues: 1-210,'G',212-220,'S',222-396,'A',398-402,'T',404-448,'A',450-695 <STR>
A;Cross-references: EMBL:X59379
R;Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sakaki, Y.
Gene 112, 189-195, 1992
A;Title: Positive and negative regulatory elements for the expression of the Alzheimer's disease amyloid precursor-encoding gene in mouse.
A;Reference number: I49485; MUID:92209998; PMID:1555768
A;Accession: I49485
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-19 <RES>
A;Cross-references: GB:D10603; NID:g220328; PIDN:BAA01456.1; PID:g220329
C;Genetics:
A;Map position: 16C3
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid; transmembrane protein

  Query Match 100.0%; Score 57; DB 2; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.0089;
  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
    ||||||||||
Db 650 KKQYTSIHHG 659

RESULT 4
S00550
Alzheimer's disease amyloid beta protein precursor - rat
N;Alternate names: beta-A4 amyloid protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C;Accession: S00550; A41245; A39820; S46251
R;Shivers, B.D.; Hilbich, C.; Multhaup, G.; Salbaum, M.; Beyreuther, K.; Seeburg, P.H.
EMBO J. 7, 1365-1370, 1988
A;Title: Alzheimer's disease amyloidogenic glycoprotein: expression pattern in rat brain suggests a role in cell contact.
A;Reference number: S00550; MUID:88312583; PMID:2900758
A;Accession: S00550

A;Molecule type: mRNA
A;Residues: 1-695 <SHI>
A;Cross-references: EMBL:X07648; NID:g55616; PIDN:CAA30488.1; PID:g55617
R;Schubert, D.; Schroeder, R.; LaCorbiere, M.; Saitoh, T.; Cole, G.
Science 241, 223-226, 1988
A;Title: Amyloid beta protein precursor is possibly a heparan sulfate proteoglycan core protein.
A;Reference number: A41245; MUID:88264430; PMID:2968652
A;Accession: A41245
A;Molecule type: protein
A;Residues: 18-37,'X',39-40,'X',42-44 <SCH>
A;Note: evidence for heparan sulfate attachment
R;Hesse, L.; Beher, D.; Masters, C.L.; Multhaup, G.
FEBS Lett. 349, 109-116, 1994
A;Title: The beta-A4 amyloid precursor protein binding to copper.
A;Reference number: S46251; MUID:94320627; PMID:7913895
A;Contents: annotation; copper binding sites
A;Note: rat peptides were isolated but not sequenced
R;Potempska, A.; Styles, J.; Mehta, P.; Kim, K.S.; Miller, D.L.
J. Biol. Chem. 266, 8464-8469, 1991
A;Title: Purification and tissue level of the beta-amyloid peptide precursor of rat brain.
A;Reference number: A39820; MUID:91217087; PMID:1673681
A;Accession: A39820
A;Status: preliminary
A;Molecule type: protein
A;Residues: 18-32 <POT>
A;Experimental source: brain
C;Comment: Deposition of amyloid protein as neurofibrillary tangles and/or plaques is characteristic of both Alzheimer's disease and Down's syndrome.
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid; glycoprotein; transmembrane protein
F;625-648/Domain: transmembrane #status predicted <TMM>

  Query Match 100.0%; Score 57; DB 2; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.0089;
  Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKQYTSIHHG 10
    ||||||||||
Db 650 KKQYTSIHHG 659

RESULT 5
QRHUA4
Alzheimer's disease amyloid beta protein precursor [validated] - human
N;Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor XIa inhibitor; proteinase nexin II (PN-II)
N;Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascular form; amyloid protein precursor splice form APP(695); amyloid protein precursor splice form APP(751); amyloid protein precursor splice form APP(770)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 28-Jul-1995 #text_change 15-Sep-2000
C;Accession: S02260; S05194; A32277; A33260; A35486; I39452; I39451; I39453; I59562; A44017; B44017; A03134; A29030; A47584; A47585; S02638; S00707; S00925; A38949; A30320; B30320; C30320; A31087; A24668; A28583; A29302; A60805; JL0038; S06121; A60355; A59011; A38384; S29076; S38252; S32539; S48148; S48692; S51186; S51185; S51184; S51183; A54238; I58075; I52250; S09010; S10737; S24127; S43644
R;Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.; Beyreuther, K.; Mueller-Hill, B.
Nucleic Acids Res. 17, 517-522, 1989
A;Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encoded by 16 exons.
A;Reference number: S02260; MUID:89128427; PMID:2783775
A;Accession: S02260
A;Molecule type: DNA
A;Residues: 1-288,'V',365-770 <LEM1>
A;Cross-references: EMBL:X13466
A;Note: alternative splice form APP(695)
R;Lemaire, H.G.
submitted to the EMBL Data Library, November 1988
A;Reference number: S05194
A;Accession: S05194
A;Molecule type: DNA
A;Residues: 1-14,'VW',17-288,'V',365-770 <LEM2>
A;Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360
A;Note: alternative splice form APP(695)
R;La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.
Biochem. Biophys. Res. Commun. 159, 297-304, 1989
A;Title: Characterization of the 5'-end region and the first two exons of the beta-protein precursor gene.
A;Reference number: A32277; MUID:89165870; PMID:2538123
A;Accession: A32277
A;Molecule type: DNA
A;Residues: 1-75 <LAF>
A;Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AAC13654.1; PID:g516074
R;Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.
Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989
A;Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows similarity to soybean trypsin inhibitor.
A;Reference number: A33260; MUID:89392030; PMID:2675837
A;Accession: A33260
A;Molecule type: DNA
A;Residues: 656-737 <JOH>
A;Cross-references: GB:M29270; NID:g178863; PIDN:AAA51768.1; PID:g178865
R;Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B.
Biochem. Biophys. Res. Commun. 170, 301-307, 1990
A;Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of hereditary cerebral hemorrhage, Dutch type: DNA and protein diagnostic assays.
A;Reference number: A35486; MUID:90321244; PMID:2196878
A;Accession: A35486
A;Molecule type: DNA
A;Residues: 672-710 <PRE1>
A;Note: 693-Gln was found in DNA isolated from HCHWA-D patients
R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 87, 257-263, 1990
A;Title: Genomic organization of the human amyloid beta-protein precursor gene.
A;Reference number: I39451; MUID:90236318; PMID:2110105
A;Accession: I39452

A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-770 <YOS1>
A;Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616
A;Accession: I39451
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-530,'QWLMPVIPAFWEAKVGR' <YOS2>
A;Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615
R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 102, 291-292, 1991
A;Reference number: A59020; MUID:91340168; PMID:1908403
A;Contents: annotation; erratum
A;Note: revised physical map for reference I39451
R;Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Duinen, S.G.; Bots, G.T.; Luyendijk, W.; Frangione, B.
Science 248, 1124-1126, 1990
A;Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type.
A;Reference number: I39453; MUID:90260663; PMID:2111584
A;Accession: I39453
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 656-737 <LEV>
A;Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620
A;Note: a mutation with 693-Gln is presented
R;Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A;Title: A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease.
A;Reference number: I59562; MUID:92022553; PMID:1925564
A;Accession: I59562
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 689-716,'F',718-737 <MUR>
A;Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721
R;Kamino, K.; Orr, H.T.; Payami, H.; Wijsman, E.M.; Alonso, M.E.; Pulst, S.M.; Anderson, L.; O'dahl, S.; Nemens, E.; White, J.A.; Sadovnick, A.D.; Ball, M.J.; Kaye, J.; Warren, A.; McInnis, M.; Antonarakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin, G.M.; Bird, T.D.; Schellenberg, G.D.
Am. J. Hum. Genet. 51, 998-1014, 1992
A;Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for the APP gene region.
A;Reference number: A44017; MUID:93035397; PMID:1415269
A;Accession: A44017
A;Molecule type: DNA
A;Residues: 687-692,'G',694-718 <KAM1>
A;Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378
A;Experimental source: familial Alzheimer disease family SB
A;Note: sequence extracted from NCBI backbone (NCBIP:115374)
A;Accession: B44017
A;Molecule type: DNA
A;Residues: 687-718 <KAM2>
A;Cross-references: GB:S45136; NID:g257379; PIDN:AAB23646.1; PID:g257380
A;Experimental source: familial Alzheimer disease family LIT
A;Note: sequence extracted from NCBI backbone (NCBIP:115376)
A;Note: this sequence has a silent mutation
R;Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.H.; Multhaup, G.; Beyreuther, K.; Muller-Hill, B.
Nature 325, 733-736, 1987
A;Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor.
A;Reference number: A03134; MUID:87144572; PMID:2881207
A;Accession: A03134
A;Molecule type: mRNA
A;Residues: 1-288,'V',365-770 <KAN>
A;Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
A;Note: alternative splice form APP(695)
R;Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A;Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascular and the neuritic plaque amyloid peptides.
A;Reference number: A29030; MUID:87231971; PMID:3035574
A;Accession: A29030
A;Molecule type: mRNA
A;Residues: 284-288,'V',365-646,'E',648-770 <ROB>
A;Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540
A;Note: the authors translated the codon GAG for residue 647 as Asp
R;Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffiotti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987
A;Title: Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease.
A;Reference number: A47584; MUID:87120328; PMID:3810169
A;Accession: A47584
A;Molecule type: mRNA
A;Residues: 674-756,'S',758-770 <GOL>
A;Cross-references: GB:M15533; NID:g178706; PIDN:AAA35540.1; PID:g178707
A;Experimental source: brain
R;Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van Keuren, M.L.; Patterson, D.; Pagan, S.; Kurnit, D.M.; Neve, R.L.
Science 235, 880-884, 1987
A;Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus.
A;Reference number: A47585; MUID:87120329; PMID:2949367
A;Accession: A47585
A;Molecule type: mRNA
A;Residues: 674-703 <TAN1>
A;Cross-references: GB:M15532; NID:g177957; PIDN:AAA51564.1; PID:g177958
R;Dyrks, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mueller-Hill, B.; Masters, C.L.; Beyreuther, K.
EMBO J. 7, 949-957, 1988
A;Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 precursor of Alzheimer's disease.
A;Reference number: S02638; MUID:88296437; PMID:2900137
A;Accession: S02638
A;Molecule type: mRNA
A;Residues: 672-678 <DYR>
R;Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; Neve, R.L.
Nature 331, 528-530, 1988

A;Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease.
A;Reference number: S00707; MUID:88122640; PMID:2893290
A;Accession: S00707
A;Molecule type: mRNA
A;Residues: 286-344,'I',365-366 <TAN2>
A;Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g929612
A;Experimental source: promyelocytic leukemia cell line HL60
A;Note: alternative splice form APP(751)
R;Ponte, P.; Gonzalez-DeWhitt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.; Davis, K.; Wallace, W.; Lieberburg, I.; Fuller, F.; Cordell, B.
Nature 331, 525-527, 1988
A;Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors.
A;Reference number: S00925; MUID:88122639; PMID:2893289
A;Accession: S00925
A;Molecule type: mRNA
A;Residues: 1-344,'I',365-770 <PO2>
A;Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721
A;Note: alternative splice form APP(751)
R;Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
Nature 331, 530-532, 1988
A;Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity.
A;Reference number: A38949; MUID:88122641; PMID:2893291
A;Accession: A38949
A;Molecule type: mRNA
A;Residues: 287-367 <KIT>
A;Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g929611
A;Experimental source: glioblastoma cell line
A;Note: alternative splice form APP(770)
R;Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ashton, R.A.; Macq, A.F.; Maloteaux, J.M.; Blume, A.J.; Octave, J.N.
Brain Res. Mol. Brain Res. 4, 121-131, 1988
A;Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three patients with sporadic Alzheimer's disease.
A;Reference number: A30320
A;Accession: A30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 284-288,'V',365-770 <VIT1>
A;Accession: B30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 122-288,'V',365-770 <VIT2>
A;Accession: C30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 606-770 <VIT3>
R;Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.; Marotta, C.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
A;Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease brain: coding and noncoding regions of the fetal precursor mRNA are expressed in the cortex.
A;Reference number: A31087; MUID:88124954; PMID:2893379
A;Accession: A31087
A;Molecule type: mRNA
A;Residues: 507-770 <ZAI>
A;Cross-references: GB:M18734; NID:g178572; PIDN:AAA51726.1; PID:g178573
A;Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 603 as Val, GTG for residue 604 as Glu, GAG for residue 605 as Leu, CTT for residue 607 as Pro, CCC for residue 608 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 655 as Ser
A;Note: the cited Genbank accession number, J03594, is not in release 101.0
R;Masters, C.L.; Multhaup, G.; Simms, G.; Pottgiesser, J.; Martins, R.N.; Beyreuther, K.

```
  Query Match            100.0%;  Score 57;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 0.01;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        725 KKQYTSIHHG 734
```

RESULT 6
JH0773
Alzheimer's disease amyloid beta protein precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 13-Aug-1999
C;Accession: JH0773
R;Okado, H.; Okamoto, H.
Biochem. Biophys. Res. Commun. 189, 1561-1568, 1992
A;Title: A Xenopus homologue of the human beta-amyloid precursor protein: developmental regulation of its gene expression.
A;Reference number: JH0773; MUID:93129227; PMID:1282805
A;Accession: JH0773
A;Molecule type: mRNA
A;Residues: 1-747 <OKA>
A;Cross-references: GB:S52417; NID:g263150; PIDN:AAB24853.1; PID:g263151
A;Experimental source: larva
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid
F;287-337/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

```
  Query Match            94.7%;  Score 54;  DB 2;  Length 747;
  Best Local Similarity  90.0%;  Pred. No. 0.036;
  Matches    9;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              |||||:||||
Db        702 KKQYTTIHHG 711
```

Search completed: October 4, 2004, 18:58:23
Job time : 14.7826 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:44:39 ; Search time 8.26087 Seconds
(without alignments)
63.032 Million cell updates/sec

Title: US-10-048-209-3
Perfect score: 57
Sequence: 1 KKQYTSIHHG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	57	100.0	737	1	A4_FUGRU	O93279 fugu rubrip
2	57	100.0	751	1	A4_SAISC	Q95241 s amyloid b
3	57	100.0	770	1	A4_CAVPO	Q60495 c amyloid b
4	57	100.0	770	1	A4_HUMAN	P05067 h amyloid b
5	57	100.0	770	1	A4_MACFA	P53601 m amyloid b
6	57	100.0	770	1	A4_MOUSE	P12023 m amyloid b
7	57	100.0	770	1	A4_PIG	P79307 s amyloid b
8	57	100.0	770	1	A4_RAT	P08592 r amyloid b
9	57	100.0	780	1	A4_TETFL	O73683 tetraodon f
10	39	68.4	452	1	NU4M_LUMTE	Q34949 lumbricus t
11	38	66.7	379	1	ISPH_SYNY3	Q55643 synechocyst
12	38	66.7	488	1	YGM9_YEAST	Q01163 saccharomyc
13	37	64.9	142	1	Y4KQ_RHISN	P55535 rhizobium s
14	36	63.2	659	1	CYOB_BUCAP	Q8k994 buchnera ap
15	36	63.2	1268	1	YRD6_CAEEL	Q09575 caenorhabdi
16	35	61.4	197	1	PSD9_CAEEL	Q10920 caenorhabdi
17	35	61.4	443	1	GAT3_HUMAN	P23771 homo sapien
18	35	61.4	443	1	GAT3_MOUSE	P23772 mus musculu
19	35	61.4	695	1	APP2_MOUSE	Q06335 mus musculu
20	35	61.4	739	1	PURL_LISIN	Q92an9 listeria in
21	35	61.4	739	1	PURL_LISMO	Q8y6c1 listeria mo
22	35	61.4	753	1	CAT2_NEUCR	Q8x182 neurospora
23	35	61.4	763	1	APP2_HUMAN	Q06481 homo sapien
24	35	61.4	765	1	APP2_RAT	P15943 rattus norv
25	34	59.6	348	1	NTRB_PROVU	P28788 proteus vul
26	34	59.6	366	1	TF3A_XENLA	P03001 xenopus lae
27	34	59.6	660	1	VNCS_PAVPN	P18547 porcine par
28	34	59.6	662	1	VNCS_PAVPK	P52502 porcine par
29	34	59.6	782	1	MUS2_STAEP	Q8cp16 staphylococ
30	34	59.6	1406	1	TOP1_CANGA	O93794 candida gla
31	33	57.9	82	1	BOF_BPP1	P22499 bacteriopha
32	33	57.9	117	1	HEMU_LINUN	P22765 lingula ung
33	33	57.9	157	1	SSRP_CHLTE	Q8kew8 chlorobium
34	33	57.9	177	1	RELX_MESAU	Q64171 mesocricetu
35	33	57.9	249	1	CREB_CHLVR	P51984 chlorohydra
36	33	57.9	326	1	ZN73_HUMAN	O43830 homo sapien
37	33	57.9	339	1	TF3A_XENBO	P17842 xenopus bor
38	33	57.9	361	1	SERC_BACHD	Q9kdm4 bacillus ha
39	33	57.9	412	1	RBA1_CAEEL	P90917 caenorhabdi
40	33	57.9	519	1	LEU1_CANBF	Q7vqj6 candidatus
41	33	57.9	530	1	UDBC_RAT	P36511 rattus norv
42	33	57.9	530	1	UDBE_RABIT	P36513 oryctolagus
43	33	57.9	580	1	SYD_MYCPE	Q8ewb7 mycoplasma
44	33	57.9	645	1	ACS2_PSEAE	Q9hv66 pseudomonas
45	33	57.9	683	1	CV19_HUMAN	Q13769 homo sapien

ALIGNMENTS

RESULT 1
A4_FUGRU
ID A4_FUGRU STANDARD; PRT; 737 AA.
AC O93279;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alzheimer's disease amyloid A4 protein homolog precursor [Contains:
DE Beta-amyloid protein (Beta-APP) (A-beta)].
GN APP.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetradontoidea; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98252138; PubMed=9599080;
RA Villard L., Tassone F., Crnogorac-Jurcevic T., Clancy K., Gardiner K.;
RT "Analysis of pufferfish homologues of the AT-rich human APP gene.";
RL Gene 210:17-24(1998).
CC -!- FUNCTION: Functional neuronal receptor which couples to
CC intracellular signaling pathway through the GTP-binding protein

```
CC       G(O) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF090120; AAD13392.1; -.
DR   HSSP; P05067; 1HZ3.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; FALSE_NEG.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
KW   Serine protease inhibitor.
FT   SIGNAL      1     18       POTENTIAL.
FT   CHAIN       19    737      ALZHEIMER'S DISEASE AMYLOID A4
FT                              PROTEIN HOMOLOG.
FT   CHAIN       639   681      BETA-AMYLOID PROTEIN (POTENTIAL).
FT   DOMAIN      19    668      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    669   689      POTENTIAL.
FT   DOMAIN      690   737      CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      286   344      BPTI/KUNITZ INHIBITOR.
FT   SITE        726   729      CLATHRIN-BINDING (BY SIMILARITY).
FT   ACT_SITE    300   301      REACTIVE BOND.
FT   DISULFID    290   340      BY SIMILARITY.
FT   DISULFID    299   323      BY SIMILARITY.
FT   DISULFID    315   336      BY SIMILARITY.
FT   CARBOHYD    522   522      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   737 AA;  82856 MW;  6FAD01E2E3B2B7E2 CRC64;

  Query Match           100.0%;  Score 57;  DB 1;  Length 737;
  Best Local Similarity  100.0%;  Pred. No. 0.0025;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        692 KKQYTSIHHG 701
```

```
RESULT 2
A4_SAISC
ID   A4_SAISC       STANDARD;      PRT;   751 AA.
AC   Q95241;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid
DE   protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha); Soluble
DE   APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42);
DE   Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-
DE   CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
GN   APP.
OS   Saimiri sciureus (Common squirrel monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX   NCBI_TaxID=9521;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Kidney, and Liver;
RX   MEDLINE=96108492; PubMed=8532114;
RA   Levy E., Amorim A., Frangione B., Walker L.C.;
RT   "Beta-amyloid precursor protein gene in squirrel monkeys with
RT   cerebral amyloid angiopathy.";
RL   Neurobiol. Aging 16:805-808(1995).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction. In vitro, copper-metallated APP induces neuronal
CC       death directly or is potentiated through Cu(II)-mediated low-
CC       density lipoprotein oxidation (By similarity). Can regulate
CC       neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
```

```
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=APP770;
CC         IsoId=Q95241-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=Q95241-2; Sequence=Not described;
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-720 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S81024; AAD14347.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Amyloid; Alternative splicing.
FT   SIGNAL       1    17       BY SIMILARITY.
FT   CHAIN       18   751       A4 PROTEIN.
FT   CHAIN       18   668       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN       18   652       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN      653   751       C99 (POTENTIAL).
FT   CHAIN      653   694       BETA-AMYLOID PROTEIN 42 (POTENTIAL).
FT   CHAIN      653   692       BETA-AMYLOID PROTEIN 40 (POTENTIAL).
FT   CHAIN      669   751       C83 (POTENTIAL).
FT   CHAIN      669   694       P3(42) (POTENTIAL).
FT   CHAIN      669   692       P3(40) (POTENTIAL).
FT   CHAIN      693   751       GAMMA-CTF(59) (POTENTIAL).
FT   CHAIN      695   751       GAMMA-CTF(57) (POTENTIAL).
FT   CHAIN      702   751       GAMMA-CTF(50) (POTENTIAL).
FT   CHAIN      721   751       C31 (POTENTIAL).
FT   DOMAIN      18   680       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   681   704       POTENTIAL.
FT   DOMAIN     705   751       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      96   110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     181   188       ZINC-BINDING (BY SIMILARITY).
```

```
FT   DOMAIN      291    341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN      316    344       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      363    428       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      504    521       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN      713    732       INTERACTION WITH G(O)-ALPHA
FT                                (BY SIMILARITY).
FT   DOMAIN      230    260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      274    280       POLY-THR.
FT   SITE        144    144       REQUIRED FOR COPPER(II) REDUCTION
FT                                (BY SIMILARITY).
FT   ACT_SITE    301    302       REACTIVE BOND.
FT   SITE        652    653       CLEAVAGE (BY BETA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        653    654       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE        668    669       CLEAVAGE (BY ALPHA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        685    685       INVOLVED IN FREE RADICAL PROPAGATION
FT                                (BY SIMILARITY).
FT   SITE        687    687       INVOLVED IN OXIDATIVE REACTIONS
FT                                (BY SIMILARITY).
FT   SITE        692    693       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                                (BY SIMILARITY).
FT   SITE        694    695       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                                (BY SIMILARITY).
FT   SITE        701    702       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT                                (BY SIMILARITY).
FT   SITE        705    715       BASOLATERAL SORTING SIGNAL
FT                                (BY SIMILARITY).
FT   SITE        720    721       CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)
FT                                (BY SIMILARITY).
FT   SITE        738    741       ENDOCYTOSIS SIGNAL.
FT   SITE        740    743       NPXY MOTIF.

  Query Match             100.0%;  Score 57;  DB 1;  Length 751;
  Best Local Similarity   100.0%;  Pred. No. 0.0026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        706 KKQYTSIHHG 715


RESULT 3
A4_CAVPO
ID   A4_CAVPO       STANDARD;      PRT;   770 AA.
AC   Q60495; Q60496;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); CTF-alpha; CTF-beta; Beta-amyloid
DE   protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); P3(42);
DE   P3(40); CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-
DE   CTF(57) (Gamma-secretase C-terminal fragment 57); C31].
GN   APP.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain, and Liver;
RX   MEDLINE=97236426; PubMed=9116031;
RA   Beck M., Mueller D., Bigl V.;
RT   "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
RT   alternative splicing.";
RL   Biochim. Biophys. Acta 1351:17-21(1997).
RN   [2]
RP   INTERACTION OF BETA-APP40 WITH APOE.
RX   MEDLINE=98007700; PubMed=9349544;
RA   Martel C.L., Mackic J.B., Matsubara E., Governale S., Miguel C.,
RA   Miao W., McComb J.G., Frangione B., Ghiso J., Zlokovic B.V.;
RT   "Isoform-specific effects of apolipoproteins E2, E3, and E4 on
RT   cerebral capillary sequestration and blood-brain barrier transport of
RT   circulating Alzheimer's amyloid beta.";
RL   J. Neurochem. 69:1995-2004(1997).
RN   [3]
RP   PROCESSING.
RX   MEDLINE=20084499; PubMed=10619481;
RA   Beck M., Brueckner M.K., Holzer M., Kaap S., Pannicke T., Arendt T.,
RA   Bigl V.;
RT   "Guinea-pig primary cell cultures provide a model to study expression
RT   and amyloidogenic processing of endogenous amyloid precursor
RT   protein.";
RL   Neuroscience 95:243-254(2000).
RN   [4]
RP   GAMMA-SECRETASE PROCESSING.
RX   MEDLINE=20576391; PubMed=11035007;
RA   Pinnix I., Musunuru U., Tun H., Sridharan A., Golde T., Eckman C.,
RA   Ziani-Cherif C., Onstead L., Sambamurti K.;
RT   "A novel gamma -secretase assay based on detection of the putative
RT   C-terminal fragment-gamma of amyloid beta protein precursor.";
RL   J. Biol. Chem. 276:481-487(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction (By similarity). In vitro, copper-metallated APP
CC       induces neuronal death directly or is potentiated through Cu(II)-
CC       mediated low-density lipoprotein oxidation (By similarity). Can
CC       regulate neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
```

```
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Beta-amyloid peptides bind to lipoproteins
CC       and apoliproteins E and J in the CSF and to HDL particles in
CC       plasma, inhibiting metal-catalyzed oxidation of lipoproteins.
CC   -!- FUNCTION: Appicans elicit adhesion of neural cells to the
CC       extracellular matrix and may regulate neurite outgrowth in the
CC       brain (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains), APPBP2 (via BaSS) and DDB1 (By similarity).
CC       Associates with microtubules in the presence of ATP and in a
CC       kinesin-dependent manner (By similarity). Soluble Abeta40 binds
CC       all three isoforms of APOE, in vitro and in vivo. When lipidated,
CC       ApoE3 appears to be the preferred amyloid binding isoform, while
CC       the apoE4 isoform-beta-APP40 complex is capable of being
CC       transported across the blood-brain barrier.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated pits
CC       (By similarity). During maturation, the immature APP (N-
CC       glycosylated in the endoplasmic reticulum) moves to the Golgi
CC       complex where complete maturation occurs (O-glycosylated and
CC       sulfated) (By similarity). After alpha-secretase cleavage, soluble
CC       APP is released into the extracellular space and the C-terminal is
CC       internalized to endosomes and lysosomes (By similarity). Some APP
CC       accumulates in secretory transport vesicles leaving the late Golgi
CC       compartment and returns to the cell surface (By similarity). APP
CC       sorts to the basolateral surface in epithelial cells (By
CC       similatity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms, missing exons 7,8 and 15, seem to
CC         exist. The L-isoforms, missing exon 15, are referred to as
CC         appicans;
CC       Name=APP770;
CC         IsoId=Q60495-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=Q60495-2; Sequence=VSP_007221, VSP_007222;
CC   -!- TISSUE SPECIFICITY: Isoform APP695 is the major isoform found in
CC       brain. The longer isoforms containing the BPTI domain are
CC       predominantly expressed in peripheral organs such as muscle and
CC       liver.
CC   -!- INDUCTION: Increased levels during neuronal differentiation.
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells.
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue (By similarity). The NPXY site is also involved in
CC       clathrin-mediated endocytosis.
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       CTF-alpha and CTF-beta. Subsequent processing of CTF-alpha by
CC       gamma-secretase yields P3 peptides. This is the major secretory
CC       pathway and is nonamyloidogenic. Alternatively,
CC       presenilin/nicastrin-mediated gamma-secretase processing of CTF-
CC       beta releases the amyloid beta proteins, amyloid-beta 40 (Abeta40)
CC       and amyloid-beta 42 (Abeta42), major components of amyloid
CC       plaques, and the corresponding cytotoxic C-terminal fragments
CC       (CTFs).
CC   -!- PTM: Proteolytically cleaved by caspase-3 during neuronal
CC       apoptosis (By similarity).
CC   -!- PTM: N- and O-glycosylated. O-linkage of chondroitin sulfate to
CC       the L-APP isoforms produces the APP proteoglycan core proteins,
CC       the appicans (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific (By similarity).
CC       Phosphorylation can affect APP processing, neuronal
CC       differentiation and interaction with other proteins.
CC   -!- PTM: Extracellular binding and reduction of copper, results in a
CC       corresponding oxidation of Cys-144 and Cys-158, and the formation
CC       of a disulfide bond (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates.
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X97631; CAA66230.1; -.
DR   EMBL; X99198; CAA67589.1; -.
DR   HSSP; P05067; 1BA4.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
```

```
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Alternative splicing; Amyloid.
FT   SIGNAL        1     17       BY SIMILARITY.
FT   CHAIN        18    770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN        18    687       SOLUBLE APP-ALPHA (BY SIMILARITY).
FT   CHAIN        18    671       SOLUBLE APP-BETA (BY SIMILARITY).
FT   CHAIN       672    770       CTF-ALPHA (BY SIMILARITY).
FT   CHAIN       672    713       BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
FT   CHAIN       672    711       BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
FT   CHAIN       688    770       CTF-BETA (BY SIMILARITY).
FT   CHAIN       688    713       P3(42) (BY SIMILARITY).
FT   CHAIN       688    711       P3(40) (BY SIMILARITY).
FT   CHAIN       712    770       GAMMA-CTF(59) (BY SIMILARITY).
FT   CHAIN       714    770       GAMMA-CTF(57) (BY SIMILARITY).

  Query Match             100.0%;  Score 57;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        725 KKQYTSIHHG 734


RESULT 4
A4_HUMAN
ID   A4_HUMAN       STANDARD;      PRT;   770 AA.
AC   P05067; P09000; P78438; Q13764; Q13778; Q13793; Q16011; Q16014;
AC   Q16019; Q16020; Q9BT38; Q9UCA9; Q9UCB6; Q9UCC8; Q9UCD1; Q9UQ58;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein) (Cerebral vascular amyloid peptide) (CVAP) (Protease
DE   nexin-II) (PN-II) (APPI) (PreA4) [Contains: Soluble APP-alpha (S-APP-
DE   alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42
DE   (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42);
DE   P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59)
DE   (Amyloid intracellular domain 59) (AID(59)); Gamma-CTF(57) (Gamma-
DE   secretase C-terminal fragment 57) (Amyloid intracellular domain 57)
DE   (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50)
DE   (Amyloid intracellular domain 50) (AID(50)); C31].
GN   APP OR A4 OR AD1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=87144572; PubMed=2881207;
RA   Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,
RA   Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;
RT   "The precursor of Alzheimer's disease amyloid A4 protein resembles a
RT   cell-surface receptor.";
RL   Nature 325:733-736(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM APP751).
RC   TISSUE=Brain;
RX   MEDLINE=88122639; PubMed=2893289;
RA   Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D.,
RA   Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,
RA   Cordell B.;
RT   "A new A4 amyloid mRNA contains a domain homologous to serine
RT   proteinase inhibitors.";
RL   Nature 331:525-527(1988).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RX   MEDLINE=89128427; PubMed=2783775;
RA   Lemaire H.-G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,
RA   Unterbeck A., Beyreuther K., Mueller-Hill B.;
RT   "The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid
RT   is encoded by 16 exons.";
RL   Nucleic Acids Res. 17:517-522(1989).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM APP770).
RX   MEDLINE=90236318; PubMed=2110105;
RA   Yoshikai S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
RT   "Genomic organization of the human amyloid beta-protein precursor
RT   gene.";
RL   Gene 87:257-263(1990).
RN   [5]
RP   ERRATUM, AND REVISIONS.
RA   Yoshikai S.-I., Sasaki H., Doh-ura K., Furuya H., Sakaki Y.;
RL   Gene 102:291-292(1991).
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM L-APP733).
RC   TISSUE=Leukocyte;
RX   MEDLINE=92268136; PubMed=1587857;
RA   Koenig G., Moenning U., Czech C., Prior R., Banati R.,
RA   Schreiter-Gasser U., Bauer J., Masters C.L., Beyreuther K.;
RT   "Identification and differential expression of a novel alternative
RT   splice isoform of the beta A4 amyloid precursor protein (APP) mRNA in
RT   leukocytes and brain microglial cells.";
RL   J. Biol. Chem. 267:10804-10809(1992).
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORM APP770).
RX   MEDLINE=97263807; PubMed=9108164;
RA   Hattori M., Tsukahara F., Furuhata Y., Tanahashi H., Hirose M.,
RA   Saito M., Tsukuni S., Sakaki Y.;
RT   "A novel method for making nested deletions and its application for
RT   sequencing of a 300 kb region of human APP locus.";
RL   Nucleic Acids Res. 25:1802-1808(1997).
RN   [8]
RP   SEQUENCE FROM N.A. (ISOFORM APP639).
RC   TISSUE=Brain;
RX   MEDLINE=22744650; PubMed=12859342;
RA   Tang K., Wang C., Shen C., Sheng S., Ravid R., Jing N.;
RT   "Identification of a novel alternative splicing isoform of human
RT   amyloid precursor protein gene, APP639.";
RL   Eur. J. Neurosci. 18:102-108(2003).
RN   [9]
```

RP SEQUENCE FROM N.A. (ISOFORM APP305).
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [10]
RP SEQUENCE OF 1-10 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016647; PubMed=3140222;
RA Schon E.A., Mita S., Sadlock J., Herbert J.;
RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
RT encodes a 95-kDa polypeptide.";
RL Nucleic Acids Res. 16:9351-9351(1988).
RN [11]
RP ERRATUM, AND REVISIONS.
RA Mita S., Sadlock J., Herbert J., Schon E.A.;
RL Nucleic Acids Res. 16:11402-11402(1988).
RN [12]
RP SEQUENCE OF 1-75 FROM N.A.
RX MEDLINE=89165870; PubMed=2538123;
RA La Fauci G., Lahiri D.K., Salton S.R., Robakis N.K.;
RT "Characterization of the 5'-end region and the first two exons of the
RT beta-protein precursor gene.";
RL Biochem. Biophys. Res. Commun. 159:297-304(1989).
RN [13]
RP SEQUENCE OF 18-50.
RC TISSUE=Fibroblast;
RX MEDLINE=87250462; PubMed=3597385;
RA van Nostrand W.E., Cunningham D.D.;
RT "Purification of protease nexin II from human fibroblasts.";
RL J. Biol. Chem. 262:8508-8514(1987).
RN [14]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP751).
RC TISSUE=Brain;
RX MEDLINE=89346754; PubMed=2569763;
RA de Sauvage F., Octave J.N.;
RT "A novel mRNA of the A4 amyloid precursor gene coding for a possibly
RT secreted protein.";
RL Science 245:651-653(1989).
RN [15]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP695).
RC TISSUE=Brain;
RX MEDLINE=87231971; PubMed=3035574;
RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
RT "Molecular cloning and characterization of a cDNA encoding the
RT cerebrovascular and the neuritic plaque amyloid peptides.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
RN [16]
RP SEQUENCE OF 286-366 FROM N.A.
RX MEDLINE=88122640; PubMed=2893290;
RA Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,
RA Gusella J.F., Neve R.L.;
RT "Protease inhibitor domain encoded by an amyloid protein precursor
RT mRNA associated with Alzheimer's disease.";
RL Nature 331:528-530(1988).
RN [17]
RP SEQUENCE OF 287-367 FROM N.A.
RX MEDLINE=88122641; PubMed=2893291;
RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
RT "Novel precursor of Alzheimer's disease amyloid protein shows
RT protease inhibitory activity.";
RL Nature 331:530-532(1988).
RN [18]
RP SEQUENCE OF 507-770 FROM N.A.
RC TISSUE=Brain cortex;
RX MEDLINE=88124954; PubMed=2893379;
RA Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
RA Marotta C.A.;
RT "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
RT disease brain: coding and noncoding regions of the fetal precursor
RT mRNA are expressed in the cortex.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
RN [19]
RP SEQUENCE OF 523-555, AND COLLAGEN-BINDING DOMAIN.
RX MEDLINE=96139497; PubMed=8576160;
RA Beher D., Hesse L., Masters C.L., Multhaup G.;
RT "Regulation of amyloid protein precursor (APP) binding to collagen and
RT mapping of the binding sites on APP and collagen type I.";
RL J. Biol. Chem. 271:1613-1620(1996).
RN [20]
RP SEQUENCE OF 655-737 FROM N.A., AND VARIANTS AD PHE-717; AD ILE-717
RP AND AD GLY-717.
RX MEDLINE=93236601; PubMed=8476439;
RA Denman R.B., Rosenzcwaig R., Miller D.L.;
RT "A system for studying the effect(s) of familial Alzheimer disease
RT mutations on the processing of the beta-amyloid peptide precursor.";
RL Biochem. Biophys. Res. Commun. 192:96-103(1993).
RN [21]
RP SEQUENCE OF 656-737 FROM N.A.
RX MEDLINE=89392030; PubMed=2675837;
RA Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
RA Little S.P.;
RT "Alzheimer's disease amyloid peptide is encoded by two exons and shows
RT similarity to soybean trypsin inhibitor.";
RL Biochem. Biophys. Res. Commun. 163:1248-1255(1989).
RN [22]

```
  Query Match             100.0%;  Score 57;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        725 KKQYTSIHHG 734


RESULT 5
A4_MACFA
ID   A4_MACFA        STANDARD;       PRT;   770 AA.
AC   P53601; Q95KN7;
DT   01-OCT-1996 (Rel. 34, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-
DE   APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);
DE   Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
GN   APP.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS APP695 AND APP770).
RC   TISSUE=Cerebellum;
RX   MEDLINE=91273117; PubMed=1905108;
RA   Podlisny M.B., Tolan D.R., Selkoe D.J.;
RT   "Homology of the amyloid beta protein precursor in monkey and human
RT   supports a primate model for beta amyloidosis in Alzheimer's
RT   disease.";
RL   Am. J. Pathol. 138:1423-1435(1991).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction. In vitro, copper-metallated APP induces neuronal
CC       death directly or is potentiated through Cu(II)-mediated low-
CC       density lipoprotein oxidation (By similarity). Can regulate
CC       neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=APP770;
CC         IsoId=P53601-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=P53601-2; Sequence=VSP_000010, VSP_000011;
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
```

```
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M58727; AAA36829.1; -.
DR   EMBL; M58726; AAA36828.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Alternative splicing; Amyloid.
FT   SIGNAL       1    17       BY SIMILARITY.
FT   CHAIN       18   770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN       18   687       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN       18   671       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN      672   770       C99 (POTENTIAL).
FT   CHAIN      672   713       BETA-AMYLOID PROTEIN 42 (POTENTIAL).
FT   CHAIN      672   711       BETA-AMYLOID PROTEIN 40 (POTENTIAL).
FT   CHAIN      688   770       C83 (POTENTIAL).
FT   CHAIN      688   713       P3(42) (POTENTIAL).
FT   CHAIN      688   711       P3(40) (POTENTIAL).
FT   CHAIN      712   770       GAMMA-CTF(59) (POTENTIAL).
FT   CHAIN      714   770       GAMMA-CTF(57) (POTENTIAL).
FT   CHAIN      721   770       GAMMA-CTF(50) (POTENTIAL).
FT   CHAIN      740   770       C31 (POTENTIAL).
FT   DOMAIN      18   699       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   700   723       POTENTIAL.
FT   DOMAIN     724   770       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      96   110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     181   188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN     291   341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN     391   423       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     491   522       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     523   540       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN     732   751       INTERACTION WITH G(O)-ALPHA
FT                              (BY SIMILARITY).
FT   DOMAIN     230   260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN     274   280       POLY-THR.
FT   SITE       144   144       REQUIRED FOR COPPER(II) REDUCTION
FT                              (BY SIMILARITY).
FT   ACT_SITE   301   302       REACTIVE BOND (BY SIMILARITY).
FT   SITE       671   672       CLEAVAGE (BY BETA-SECRETASE)
FT                              (BY SIMILARITY).
FT   SITE       672   673       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE       687   688       CLEAVAGE (BY ALPHA-SECRETASE)
FT                              (BY SIMILARITY).
FT   SITE       704   704       IMPLICATED IN FREE RADICAL PROPAGATION
FT                              (BY SIMILARITY).
FT   SITE       706   706       INVOLVED IN OXIDATIVE REACTIONS
FT                              (BY SIMILARITY).
FT   SITE       711   712       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                              (BY SIMILARITY).
FT   SITE       713   714       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                              (BY SIMILARITY).
FT   SITE       720   721       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT                              (BY SIMILARITY).
FT   SITE       724   734       BASOLATERAL SORTING SIGNAL
FT                              (BY SIMILARITY).
FT   SITE       739   740       CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)

  Query Match           100.0%;  Score 57;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        725 KKQYTSIHHG 734


RESULT 6
A4_MOUSE
ID   A4_MOUSE        STANDARD;      PRT;   770 AA.
AC   P12023; P97487; P97942; Q99K32;
DT   01-OCT-1989 (Rel. 12, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
```

```
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains:
DE   Soluble APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99
DE   (APP-C99); Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein
DE   40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase
DE   C-terminal fragment 59) (Amyloid intracellular domain 59) (AID(59))
DE   (APP-C59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57)
DE   (Amyloid intracellular domain 57) (AID(57)) (APP-C57); Gamma-CTF(50)
DE   (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain
DE   50) (AID(50)); C31].
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=88106489; PubMed=3322280;
RA   Yamada T., Sasaki H., Furuya H., Miyata T., Goto I., Sakaki Y.;
RT   "Complementary DNA for the mouse homolog of the human amyloid beta
RT   protein precursor.";
RL   Biochem. Biophys. Res. Commun. 149:665-671(1987).
RN   [2]
RP   REVISIONS.
RA   Yamada T.;
RL   Submitted (MAR-1988) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=92096458; PubMed=1756177;
RA   de Strooper B., van Leuven F., van den Berghe H.;
RT   "The amyloid beta protein precursor or proteinase nexin II from mouse
RT   is closer related to its human homolog than previously reported.";
RL   Biochim. Biophys. Acta 1129:141-143(1991).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   STRAIN=SAMP8; TISSUE=Hippocampus;
RX   MEDLINE=21130647; PubMed=11235921;
RA   Kumar V.B., Vyas K., Franko M., Choudhary V., Buddhiraju C.,
RA   Alvarez J., Morley J.E.;
RT   "Molecular cloning, expression, and regulation of hippocampal amyloid
RT   precursor protein of senescence accelerated mouse (SAMP8).";
RL   Biochem. Cell Biol. 79:57-67(2001).
RN   [5]
RP   SEQUENCE OF 1-19 FROM N.A.
RX   MEDLINE=92209998; PubMed=1555768;
RA   Izumi R., Yamada T., Yoshikai S.I., Sasaki H., Hattori M.,
RA   Sakai Y.;
RT   "Positive and negative regulatory elements for the expression of the
RT   Alzheimer's disease amyloid precursor-encoding gene in mouse.";
RL   Gene 112:189-195(1992).
RN   [6]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORM APP770).
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain, and Kidney;
RX   MEDLINE=89149813; PubMed=2493250;
RA   Yamada T., Sasaki H., Dohura K., Goto I., Sakaki Y.;
RT   "Structure and expression of the alternatively-spliced forms of mRNA
RT   for the mouse homolog of Alzheimer's disease amyloid beta protein
RT   precursor.";
RL   Biochem. Biophys. Res. Commun. 158:906-912(1989).
RN   [8]
RP   SEQUENCE OF 289-364 FROM N.A.
RC   STRAIN=CD-1; TISSUE=Placenta;
RX   MEDLINE=89345111; PubMed=2569710;
RA   Fukuchi K., Martin G.M., Deeb S.S.;
RT   "Sequence of the protease inhibitor domain of the A4 amyloid protein
RT   precursor of Mus domesticus.";
RL   Nucleic Acids Res. 17:5396-5396(1989).
RN   [9]
RP   SEQUENCE OF 656-737 FROM N.A.
RC   STRAIN=129/Sv;
RA   Wragg M.A., Busfield F., Duff K., Korenblat K., Capecchi M.,
RA   Loring J.F., Goate A.M.;
RT   "Introduction of six mutations into the mouse genome using 'Hit and
RT   Run' gene-targeting: introduction of familial Alzheimer's disease
RT   mutations into the mouse amyloid precursor protein gene and
RT   humanization of the A-beta fragment.";
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN   [10]
RP   TISSUE SPECIFICITY OF ALTERNATIVE SPLICED FORMS.
RX   MEDLINE=93287808; PubMed=8510506;
RA   Sola C., Mengod G., Ghetti B., Palacios J.M., Triarhou L.C.;
RT   "Regional distribution of the alternatively spliced isoforms of beta
RT   APP RNA transcript in the brain of normal, heterozygous and
RT   homozygous weaver mutant mice as revealed by in situ hybridization
RT   histochemistry.";
RL   Brain Res. Mol. Brain Res. 17:340-346(1993).
RN   [11]
```

```
RP   INTERACTION WITH KNS2.
RX   MEDLINE=21010507; PubMed=11144355;
RA   Kamal A., Stokin G.B., Yang Z., Xia C.-H., Goldstein L.S.;
RT   "Axonal transport of amyloid precursor protein is mediated by direct
RT   binding to the kinesin light chain subunit of kinesin-I.";
RL   Neuron 28:449-459(2000).
RN   [12]
RP   C-TERMINAL PROTEIN-PROTEIN INTERACTIONS, AND MUTAGENESIS OF TYR-728;
RP   THR-743; TYR-757; ASN-759 AND TYR-762.
RX   MEDLINE=21408156; PubMed=11517249;
RA   Matsuda S., Yasukawa T., Homma Y., Ito Y., Niikura T., Hiraki T.,
RA   Hirai S., Ohno S., Kita Y., Kawasumi M., Kouyama K., Yamamoto T.,
RA   Kyriakis J.M., Nishimoto I.;
RT   "C-jun N-terminal kinase (JNK)-interacting protein-1b/islet-brain-1
RT   scaffolds Alzheimer's amyloid precursor protein with JNK.";
RL   J. Neurosci. 21:6597-6607(2001).
RN   [13]
RP   INTERACTION WITH MAPK8IP1, AND PHOSPHORYLATION.
RX   MEDLINE=22028091; PubMed=11912189;
RA   Taru H., Iijima K.-I., Hase M., Kirino Y., Yagi Y., Suzuki T.;
RT   "Interaction of Alzheimer's beta-amyloid precursor family proteins
RT   with scaffold proteins of the JNK signaling cascade.";
RL   J. Biol. Chem. 277:20070-20078(2002).
RN   [14]
RP   INTERACTION OF CTF PEPTIDES WITH NUMB.
RX   MEDLINE=22008109; PubMed=12011466;
RA   Roncarati R., Sestan N., Scheinfeld M.H., Berechid B.E., Lopez P.A.,
RA   Meucci O., McGlade J.C., Rakic P., D'Adamio L.;
RT   "The gamma-secretase-generated intracellular domain of beta-amyloid
RT   precursor protein binds Numb and inhibits Notch signaling.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:7102-7107(2002).
RN   [15]
RP   GAMMA-SECRETASE PROCESSING, AND INTERACTION WITH APBB1.
RX   MEDLINE=21437805; PubMed=11553691;
RA   Cupers P., Orlans I., Craessaerts K., Annaert W., De Strooper B.;
RT   "The amyloid precursor protein (APP)-cytoplasmic fragment generated by
RT   gamma-secretase is rapidly degraded but distributes partially in a
RT   nuclear fraction of neurones in culture.";
RL   J. Neurochem. 78:1168-1178(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions. Can promote transcription activation through binding
CC       to APBB1/Tip60 and inhibit Notch signaling through interaction
CC       with Numb. Couples to apoptosis-inducing pathways such as those
CC       mediated by G(O) and JIP. Inhibits G(O) alpha ATPase activity (By
CC       similarity). Acts as a kinesin I membrane receptor, mediating the
CC       axonal transport of beta-secretase and presenilin 1. May be
CC       involved in copper homeostasis/oxidative stress through copper ion
CC       reduction. Can regulate neurite outgrowth through binding to
CC       components of the extracellular matrix such as heparin and
CC       collagen I and IV (By similarity). The splice isoforms that
CC       contain the BPTI domain possess protease inhibitor activity (By
CC       similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
CC       only weakly transient metals and have little reducing activity due
CC       to substitutions of transient metal chelating residues. Beta-APP42
CC       may activate mononuclear phagocytes in the brain and elicit
CC       inflammatory responses. Promotes both tau aggregation and TPK II-
CC       mediated phosphorylation (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis.
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1, Numb and Dab1. Binding to Dab1 inhibits
CC       its serine phosphorylation. Also interacts with GPCR-like protein
CC       BPP, FPRL1, APPBP1, IB1, KNS2 (via its TPR domains), APPBP2 (via
CC       BaSS) and DDB1 (By similarity). In vitro, it binds MAPT via the
CC       MT-binding domains (By similarity). Associates with microtubules
CC       in the presence of ATP and in a kinesin-dependent manner (By
CC       similarity). Interacts, through a C-terminal domain, with GNAO1
CC       (By similarity). Amyloid beta-42 binds CHRNA7 in hippocampal
CC       neurons (By similarity). Beta-amyloid associates with HADH2 (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
```

```
  Query Match           100.0%;  Score 57;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        725 KKQYTSIHHG 734
```

```
RESULT 7
A4_PIG
ID   A4_PIG          STANDARD;      PRT;   770 AA.
AC   P79307; Q29023; Q9TUI0;
DT   01-NOV-1997 (Rel. 35, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-
DE   APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);
DE   Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kimura A., Takahashi T.;
RT   "Amyloid precursor protein 770.";
```

```
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-136 FROM N.A.
RC   TISSUE=Small intestine;
RA   Winteroe A.K., Fredholm M.;
RT   "Evaluation and characterization of a porcine small intestine cDNA
RT   library.";
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 667-723 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=92017079; PubMed=1656157;
RA   Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT   "Conservation of the sequence of the Alzheimer's disease amyloid
RT   peptide in dog, polar bear and five other mammals by cross-species
RT   polymerase chain reaction analysis.";
RL   Brain Res. Mol. Brain Res. 10:299-305(1991).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction (By similarity). In vitro, copper-metallated APP
CC       induces neuronal death directly or is potentiated through Cu(II)-
CC       mediated low-density lipoprotein oxidation (By similarity). Can
CC       regulate neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- PTM: Extracellular binding and reduction of copper, results in a
CC       corresponding oxidation of Cys-144 and Cys-158, and the formation
CC       of a disulfide bond (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
```

```
DR   EMBL; AB032550; BAA84580.1; -.
DR   EMBL; Z84022; CAB06313.1; -.
DR   EMBL; X56127; CAA39592.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Amyloid.
FT   SIGNAL       1    17       BY SIMILARITY.
FT   CHAIN       18   770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN       18   687       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN       18   671       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN      672   770       C99 (BY SIMILARITY).
FT   CHAIN      672   713       BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
FT   CHAIN      672   711       BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
FT   CHAIN      688   770       C83 (BY SIMILARITY).
FT   CHAIN      688   713       P3(42) (BY SIMILARITY).
FT   CHAIN      688   711       P3(40) (BY SIMILARITY).
FT   CHAIN      712   770       GAMMA-CTF(59).
FT   CHAIN      714   770       GAMMA-CTF(57).
FT   CHAIN      721   770       GAMMA-CTF(50) (BY SIMILARITY).
FT   CHAIN      740   770       C31 (DURING APOPTOSIS) (BY SIMILARITY).
FT   DOMAIN      18   699       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   700   723       POTENTIAL.
FT   DOMAIN     724   770       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      96   110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     135   155       COPPER-BINDING (BY SIMILARITY).
FT   DOMAIN     181   188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN     291   341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN     391   423       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     491   522       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     523   540       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN     732   751       INTERACTION WITH G(O)-ALPHA (BY
FT                              SIMILARITY).
FT   DOMAIN     230   260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN     274   280       POLY-THR.
FT   SITE       144   144       REQUIRED FOR COPPER(II) REDUCTION
FT                              (BY SIMILARITY).
FT   ACT_SITE   301   302       REACTIVE BOND (BY SIMILARITY).
FT   SITE       671   672       CLEAVAGE (BY BETA-SECRETASE)
FT                              (BY SIMILARITY).
FT   SITE       672   673       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE       687   688       CLEAVAGE (BY ALPHA-SECRETASE)
FT                              (BY SIMILARITY).
FT   SITE       704   704       IMPLICATED IN FREE RADICAL PROPAGATION
FT                              (BY SIMILARITY).
FT   SITE       706   706       INVOLVED IN OXIDATIVE REACTIONS
FT                              (BY SIMILARITY).
FT   SITE       711   712       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                              (BY SIMILARITY).
FT   SITE       713   714       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                              (BY SIMILARITY).
FT   SITE       720   721       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)

  Query Match             100.0%;  Score 57;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        725 KKQYTSIHHG 734


RESULT 8
A4_RAT
ID   A4_RAT          STANDARD;      PRT;   770 AA.
AC   P08592;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid
DE   protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains: Soluble
DE   APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99; Beta-
DE   amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40);
DE   C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal
DE   fragment 59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57);
DE   Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50); C31].
GN   APP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=88312583; PubMed=2900758;
RA   Shivers B.D., Hilbich C., Multhaup G., Salbaum J.M., Beyreuther K.,
RA   Seeburg P.H.;
RT   "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
RT   in rat brain suggests a role in cell contact.";
RL   EMBO J. 7:1365-1370(1988).
RN   [2]
RP   SEQUENCE OF 289-364 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=89183625; PubMed=2648331;
RA   Kang J., Mueller-Hill B.;
RT   "The sequence of the two extra exons in rat preA4.";
RL   Nucleic Acids Res. 17:2130-2130(1989).
RN   [3]
RP   SEQUENCE OF 720-730, AND MASS SPECTROMETRY.
RX   MEDLINE=21443797; PubMed=11483588;
```

RA Gu Y., Misonou H., Sato T., Dohmae N., Takio K., Ihara Y.;
RT "Distinct intramembrane cleavage of the beta-amyloid precursor protein
RT family resembling gamma-secretase-like cleavage of Notch.";
RL J. Biol. Chem. 276:35235-35238(2001).
RN [4]
RP ALTERNATIVE SPLICING.
RX MEDLINE=96187032; PubMed=8624099;
RA Sandbrink R., Masters C.L., Beyreuther K.;
RT "APP gene family. Alternative splicing generates functionally related
RT isoforms.";
RL Ann. N.Y. Acad. Sci. 777:281-287(1996).
RN [5]
RP TISSUE SPECIFICITY OF APPICAN.
RX MEDLINE=95263526; PubMed=7744833;
RA Shioi J., Pangalos M.N., Ripellino J.A., Vassilacopoulou D.,
RA Mytilineou C., Margolis R.U., Robakis N.K.;
RT "The Alzheimer amyloid precursor proteoglycan (appican) is present in
RT brain and is produced by astrocytes but not by neurons in primary
RT neural cultures.";
RL J. Biol. Chem. 270:11839-11844(1995).
RN [6]
RP TISSUE SPECIFICITY OF ISOFORMS.
RX MEDLINE=97150061; PubMed=8996834;
RA Sandbrink R., Monning U., Masters C.L., Beyreuther K.;
RT "Expression of the APP gene family in brain cells, brain development
RT and aging.";
RL Gerontology 43:119-131(1997).
RN [7]
RP INTERACTION WITH DDB1, AND MUTAGENESIS OF TYR-757; ASN-759 AND
RP TYR-762.
RX MEDLINE=99127916; PubMed=9930726;
RA Watanabe T., Sukegawa J., Tomita S., Iijima K.-I., Oguchi S.,
RA Suzuki T., Nairn A.C., Greengard P.;
RT "A 127-kDa protein (UV-DDB) binds to the cytoplasmic domain of the
RT Alzheimer's amyloid precursor protein.";
RL J. Neurochem. 72:549-556(1999).
RN [8]
RP INTERACTION WITH GNAO1, AND MUTAGENESIS OF 732-HIS-HIS-733.
RX MEDLINE=99162676; PubMed=10024358;
RA Brouillet E., Trembleau A., Galanaud D., Volovitch M., Bouillot C.,
RA Valenza C., Prochiantz A., Allinquant B.;
RT "The amyloid precursor protein interacts with Go heterotrimeric
RT protein within a cell compartment specialized in signal
RT transduction.";
RL J. Neurosci. 19:1717-1727(1999).
RN [9]
RP CHARACTERISTICS OF APPICAN, AND MUTAGENESIS OF SER-656.
RX MEDLINE=95256193; PubMed=7737970;
RA Pangalos M.N., Efthimiopoulos S., Shioi J., Robakis N.K.;
RT "The chondroitin sulfate attachment site of appican is formed by
RT splicing out exon 15 of the amyloid precursor gene.";
RL J. Biol. Chem. 270:10388-10391(1995).
RN [10]
RP BETA-AMYLOID METAL-BINDING.
RX MEDLINE=99316162; PubMed=10386999;
RA Huang X., Atwood C.S., Hartshorn M.A., Multhaup G., Goldstein L.E.,
RA Scarpa R.C., Cuajungco M.P., Gray D.N., Lim J., Moir R.D., Tanzi R.E.,
RA Bush A.I.;
RT "The A beta peptide of Alzheimer's disease directly produces hydrogen
RT peroxide through metal ion reduction.";
RL Biochemistry 38:7609-7616(1999).
RN [11]
RP BETA-AMYLOID ZINC BINDING.
RX MEDLINE=99343552; PubMed=10413512;
RA Liu S.T., Howlett G., Barrow C.J.;
RT "Histidine-13 is a crucial residue in the zinc ion-induced aggregation
RT of the A beta peptide of Alzheimer's disease.";
RL Biochemistry 38:9373-9378(1999).
RN [12]
RP IMPORTANCE OF GLY-704 IN FREE RADICAL PROPAGATION, AND MUTAGENESIS OF
RP GLY-704.
RX MEDLINE=21956095; PubMed=11959460;
RA Kanski J., Varadarajan S., Aksenova M., Butterfield D.A.;
RT "Role of glycine-33 and methionine-35 in Alzheimer's amyloid beta-
RT peptide 1-42-associated oxidative stress and neurotoxicity.";
RL Biochim. Biophys. Acta 1586:190-198(2001).
RN [13]
RP PHOSPHORYLATION.
RX MEDLINE=97239592; PubMed=9085254;
RA Oishi M., Nairn A.C., Czernik A.J., Lim G.S., Isohara T., Gandy S.E.,
RA Greengard P., Suzuki T.;
RT "The cytoplasmic domain of Alzheimer's amyloid precursor protein is
RT phosphorylated at Thr654, Ser655, and Thr668 in adult rat brain and
RT cultured cells.";
RL Mol. Med. 3:111-123(1997).
RN [14]
RP PHOSPHORYLATION ON SER-730.
RX MEDLINE=99262094; PubMed=10329382;
RA Isohara T., Horiuchi A., Watanabe T., Ando K., Czernik A.J., Uno I.,
RA Greengard P., Nairn A.C., Suzuki T.;
RT "Phosphorylation of the cytoplasmic domain of Alzheimer's beta-amyloid
RT precursor protein at Ser655 by a novel protein kinase.";
RL Biochem. Biophys. Res. Commun. 258:300-305(1999).
RN [15]
RP PHOSPHORYLATION, INDUCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
RP THR-743.
RX MEDLINE=99274744; PubMed=10341243;
RA Ando K., Oishi M., Takeda S., Iijima K.-I., Isohara T., Nairn A.C.,
RA Kirino Y., Greengard P., Suzuki T.;
RT "Role of phosphorylation of Alzheimer's amyloid precursor protein
RT during neuronal differentiation.";
RL J. Neurosci. 19:4421-4427(1999).
RN [16]
RP PHOSPHORYLATION ON THR-743.
RX MEDLINE=20396183; PubMed=10936190;
RA Iijima K.-I., Ando K., Takeda S., Satoh Y., Seki T., Itohara S.,
RA Greengard P., Kirino Y., Nairn A.C., Suzuki T.;
RT "Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor
RT protein by cyclin-dependent kinase 5.";
RL J. Neurochem. 75:1085-1091(2000).
RN [17]
RP CARBOHYDRATE STRUCTURE OF APPICAN.
RX MEDLINE=21463085; PubMed=11479316;

```
RA   Tsuchida K., Shioi J., Yamada S., Boghosian G., Wu A., Cai H.,
RA   Sugahara K., Robakis N.K.;
RT   "Appican, the proteoglycan form of the amyloid precursor protein,
RT   contains chondroitin sulfate E in the repeating disaccharide region
RT   and 4-O-sulfated galactose in the linkage region.";
RL   J. Biol. Chem. 276:37155-37160(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP. Inhibits
CC       G(O) alpha ATPase activity. Acts as a kinesin I membrane receptor,
CC       mediating the axonal transport of beta-secretase and presenilin 1
CC       (By similarity). May be involved in copper homeostasis/oxidative
CC       stress through copper ion reduction. Can regulate neurite
CC       outgrowth through binding to components of the extracellular
CC       matrix such as heparin and collagen I and IV (By similarity). The
CC       splice isoforms that contain the BPTI domain possess protease
CC       inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
CC       only weakly transient metals and have little reducing activity due
CC       to substitutions of transient metal chelating residues. Beta-APP42
CC       may activate mononuclear phagocytes in the brain and elicit
CC       inflammatory responses. Promotes both tau aggregation and TPK II-
CC       mediated phosphorylation (By similarity).
CC   -!- FUNCTION: Appicans elicit adhesion of neural cells to the
CC       extracellular matrix and may regulate neurite outgrowth in the
CC       brain.
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains), APPBP2 (via BaSS) (By similarity) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity). Interacts,
CC       through a C-terminal domain, with GNAO1. Amyloid beta-42 binds
CC       CHRNA7 in hippocampal neurons (By similarity). Beta-amyloid
CC       associates with HADH2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the

  Query Match           100.0%;  Score 57;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
                |||||||||| 
Db        725 KKQYTSIHHG 734


RESULT 9
A4_TETFL
ID   A4_TETFL       STANDARD;      PRT;   780 AA.
AC   073683;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Alzheimer's disease amyloid A4 protein homolog precursor [Contains:
DE   Beta-amyloid protein (Beta-APP) (A-beta)].
GN   APP.
OS   Tetraodon fluviatilis (Puffer fish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=47145;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98252138; PubMed=9599080;
RA   Villard L., Tassone F., Crnogorac-Jurcevic T., Clancy K., Gardiner K.;
RT   "Analysis of pufferfish homologues of the AT-rich human APP gene.";
RL   Gene 210:17-24(1998).
CC   -!- FUNCTION: Functional neuronal receptor which couples to
CC       intracellular signaling pathway through the GTP-binding protein
CC       G(O) (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AF018165; AAC41275.1; -.
DR   HSSP; P05067; 1HZ3.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
```

```
DR   PROSITE; PS00280; BPTI_KUNITZ_1; FALSE_NEG.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Glycoprotein; Amyloid; Neurone; Transmembrane; Signal;
KW   Serine protease inhibitor.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    780       ALZHEIMER'S DISEASE AMYLOID A4 PROTEIN
FT                                 HOMOLOG.
FT   CHAIN       682    724       BETA-AMYLOID PROTEIN (POTENTIAL).
FT   DOMAIN       19    711       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    712    732       POTENTIAL.
FT   DOMAIN      733    780       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      323    382       BPTI/KUNITZ INHIBITOR.
FT   SITE        769    772       CLATHRIN-BINDING (BY SIMILARITY).
FT   DISULFID    327    378       BY SIMILARITY.
FT   DISULFID    336    361       BY SIMILARITY.
FT   CARBOHYD    560    560       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   780 AA;  88238 MW;  60071BE945201910 CRC64;

  Query Match             100.0%;  Score 57;  DB 1;  Length 780;
  Best Local Similarity   100.0%;  Pred. No. 0.0027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        735 KKQYTSIHHG 744


RESULT 10
NU4M_LUMTE
ID   NU4M_LUMTE     STANDARD;      PRT;   452 AA.
AC   Q34949;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   NADH-ubiquinone oxidoreductase chain 4 (EC 1.6.5.3).
GN   ND4.
OS   Lumbricus terrestris (Common earthworm).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC   Lumbricina; Lumbricidae; Lumbricus.
OX   NCBI_TaxID=6398;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96042914; PubMed=8536978;
RA   Boore J.L., Brown W.M.;
RT   "Complete sequence of the mitochondrial DNA of the annelid worm
RT   Lumbricus terrestris.";
RL   Genetics 141:305-319(1995).
CC   -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U24570; AAC46873.1; -.
DR   PIR; S58994; S58994.
DR   InterPro; IPR003918; NADHub_oxred4.
DR   InterPro; IPR003916; NADHub_oxred5.
DR   InterPro; IPR001750; Oxidored_q1.
DR   InterPro; IPR000260; Oxidored_q5_N.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF01059; oxidored_q5_N; 1.
DR   PRINTS; PR01434; NADHDHGNASE5.
DR   PRINTS; PR01437; NUOXDRDTASE4.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
SQ   SEQUENCE   452 AA;  50482 MW;  604CF0F63129DEDE CRC64;

  Query Match             68.4%;  Score 39;  DB 1;  Length 452;
  Best Local Similarity   85.7%;  Pred. No. 4.8;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 YTSIHHG 10
              |||:|||
Db        405 YTSMHHG 411


RESULT 11
ISPH_SYNY3
ID   ISPH_SYNY3     STANDARD;      PRT;   379 AA.
AC   Q55643; Q55066;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   4-hydroxy-3-methylbut-2-enyl diphosphate reductase (EC 1.17.1.2).
GN   ISPH OR LYTB OR SLR0348.
OS   Synechocystis sp. (strain PCC 6803).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX   NCBI_TaxID=1148;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97449302; PubMed=9305771;
RA   Cassier-Chauvat C., Poncelet M., Chauvat F.;
RT   "Three insertion sequences from the cyanobacterium Synechocystis
RT   PCC6803 support the occurrence of horizontal DNA transfer among
RT   bacteria.";
RL   Gene 195:257-266(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127529; PubMed=8590279;
RA   Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA   Sugiura M., Tabata S.;
RT   "Sequence analysis of the genome of the unicellular cyanobacterium
RT   Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT   region from map positions 64% to 92% of the genome.";
RL   DNA Res. 2:153-166(1995).
RN   [3]
RP   PATHWAY.
RX   MEDLINE=20461233; PubMed=11004185;
RA   Cunningham F.X. Jr., Lafond T.P., Gantt E.;
RT   "Evidence of a role for LytB in the nonmevalonate pathway of
```

```
RT   isoprenoid biosynthesis.";
RL   J. Bacteriol. 182:5841-5848(2000).
CC   -!- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
CC       into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate
CC       (DMAPP).
CC   -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate + NAD(P)(+) + H(2)O =
CC       (E)-4-hydroxy-3-methylbut-2-en-1-yl diphosphate + NAD(P)H.
CC   -!- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh
CC       (last) step.
CC   -!- SIMILARITY: Belongs to the ispH family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U38915; AAB72119.1; -.
DR   EMBL; D64000; BAA10159.1; ALT_INIT.
DR   HAMAP; MF_00191; -; 1.
DR   InterPro; IPR003451; LytB.
DR   Pfam; PF02401; LYTB; 1.
DR   TIGRFAMs; TIGR00216; ispH_lytB; 1.
KW   Isoprene biosynthesis; Complete proteome; Oxidoreductase; NADP.
SQ   SEQUENCE   379 AA;  42415 MW;  4CD0167AF5D4957E CRC64;

  Query Match             66.7%;  Score 38;  DB 1;  Length 379;
  Best Local Similarity   70.0%;  Pred. No. 6.2;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||::||| ||
Db        145 KKEHTSIIHG 154


RESULT 12
YGM9_YEAST
ID   YGM9_YEAST     STANDARD;      PRT;   488 AA.
AC   Q01163;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical 55.6 kDa protein in CEG1-SOH1 intergenic region.
GN   YGL129C OR G2856.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96287651; PubMed=8686378;
RA   Rodriguez-Belmonte E., Rodriguez Torres A.M., Tizon B., Cadahia J.L.,
RA   Gonzalez-Siso I., Ramil E., Becerra M., Gonzalez-Dominguez M.,
RA   Cerdan E.;
RT   "Sequence analysis of a 10 kb DNA fragment from yeast chromosome VII
RT   reveals a novel member of the DnaJ family.";
RL   Yeast 12:145-148(1996).
RN   [2]
RP   SEQUENCE OF 39-488 FROM N.A.
RC   STRAIN=DBY939;
RX   MEDLINE=92250598; PubMed=1315757;
RA   Shibagaki Y., Itoh N., Yamada H., Nagata S., Mizumoto K.;
RT   "mRNA capping enzyme. Isolation and characterization of the gene
RT   encoding mRNA guanylytransferase subunit from Saccharomyces
RT   cerevisiae.";
RL   J. Biol. Chem. 267:9521-9528(1992).
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X87252; CAA60702.1; -.
DR   EMBL; Z72651; CAA96838.1; -.
DR   EMBL; D10263; BAA01104.1; -.
DR   PIR; S64140; S64140.
DR   GermOnline; 141177; -.
DR   SGD; S0003097; RSM23.
DR   GO; GO:0005763; C:mitochondrial small ribosomal subunit; IDA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IDA.
DR   GO; GO:0006412; P:protein biosynthesis; IDA.
KW   Hypothetical protein; ATP-binding.
FT   NP_BIND     188    195       ATP (POTENTIAL).
SQ   SEQUENCE   488 AA;  55562 MW;  C59E8AA8858BEA18 CRC64;

  Query Match             66.7%;  Score 38;  DB 1;  Length 488;
  Best Local Similarity   75.0%;  Pred. No. 8.1;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 KQYTSIHH 9
              ||| |:||
Db        137 KQYKSLHH 144


RESULT 13
Y4KQ_RHISN
ID   Y4KQ_RHISN     STANDARD;      PRT;   142 AA.
AC   P55535;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   Very hypothetical 15.3 kDa protein Y4KQ.
GN   Y4KQ.
OS   Rhizobium sp. (strain NGR234).
OG   Plasmid sym pNGR234a.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX   NCBI_TaxID=394;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97305956; PubMed=9163424;
RA   Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA   Perret X.;
RT   "Molecular basis of symbiosis between Rhizobium and legumes.";
RL   Nature 387:394-401(1997).
CC   -!- SIMILARITY: NONE OBVIOUS. COULD BE A FRAGMENT.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000082; AAB91748.1; -.
DR   PIR; T10871; T10871.
KW   Hypothetical protein; Plasmid.
SQ   SEQUENCE   142 AA;  15332 MW;  4277C6BDD33E8DF1 CRC64;

  Query Match             64.9%;  Score 37;  DB 1;  Length 142;
  Best Local Similarity   60.0%;  Pred. No. 3.3;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              |:  ||:|||
Db        127 KRSGTSLHHG 136


RESULT 14
CYOB_BUCAP
ID   CYOB_BUCAP     STANDARD;      PRT;   659 AA.
AC   Q8K994;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ubiquinol oxidase polypeptide I (EC 1.10.3.-) (Cytochrome O subunit 1)
DE   (Oxidase BO(3) subunit 1) (Cytochrome O ubiquinol oxidase subunit 1).
GN   CYOB OR BUSG455.
OS   Buchnera aphidicola (subsp. Schizaphis graminum).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Buchnera.
OX   NCBI_TaxID=98794;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22084549; PubMed=12089438;
RA   Tamas I., Klasson L., Canbaeck B., Naeslund A.K., Eriksson A.-S.,
RA   Wernegreen J.J., Sandstroem J.P., Moran N.A., Andersson S.G.E.;
RT   "50 million years of genomic stasis in endosymbiotic bacteria.";
RL   Science 296:2376-2379(2002).
CC   -!- FUNCTION: Cytochrome O terminal oxidase complex is the component
CC       of the aerobic respiratory chain that predominates when cells are
CC       grown at high aeration. This ubiquinol oxidase shows proton pump
CC       activity across the membrane in addition to the electron transfer
CC       (By similarity).
CC   -!- CATALYTIC ACTIVITY: Ubiquinol-8 + O(2) = Ubiquinone-8 + H(2)O.
CC   -!- COFACTOR: Contains two protoheme IX (heme B55 and B562) and
CC       copper B (By similarity).
CC   -!- PATHWAY: Ubiquinol oxidase catalyzes the terminal step in the
CC       electron transport chain.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC   -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE014121; AAM67998.1; -.
DR   InterPro; IPR000883; COX1.
DR   Pfam; PF00115; COX1; 1.
DR   PRINTS; PR01165; CYCOXIDASEI.
DR   PROSITE; PS00077; COX1; 1.
KW   Oxidoreductase; Respiratory chain; Transmembrane; Heme; Copper;
KW   Hydrogen ion transport; Complete proteome.
FT   DOMAIN       1     14       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     15    35       POTENTIAL.
FT   DOMAIN       36    58       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     59    79       POTENTIAL.
FT   DOMAIN       80    106      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     107   127      POTENTIAL.
FT   DOMAIN       128   145      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     146   166      POTENTIAL.
FT   DOMAIN       167   189      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     190   210      POTENTIAL.
FT   DOMAIN       211   225      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     226   246      POTENTIAL.
FT   DOMAIN       247   277      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     278   298      POTENTIAL.
FT   DOMAIN       299   309      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     310   330      POTENTIAL.
FT   DOMAIN       331   347      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     348   368      POTENTIAL.
FT   DOMAIN       369   380      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     381   401      POTENTIAL.
FT   DOMAIN       402   413      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     414   434      POTENTIAL.
FT   DOMAIN       435   456      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     457   477      POTENTIAL.
FT   DOMAIN       478   490      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     491   511      POTENTIAL.
FT   DOMAIN       512   580      CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     581   601      POTENTIAL.
FT   DOMAIN       602   605      EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM     606   626      POTENTIAL.
FT   DOMAIN       627   659      CYTOPLASMIC (POTENTIAL).
FT   METAL        106   106      IRON (HEME B AXIAL LIGAND) (PROBABLE).
FT   METAL        284   284      COPPER B (PROBABLE).
FT   METAL        288   288      COPPER B (PROBABLE).
FT   METAL        333   333      COPPER B (PROBABLE).
```

```
FT   METAL       334    334       COPPER B (PROBABLE).
FT   METAL       419    419       IRON (HEME O AXIAL LIGAND) (PROBABLE).
FT   METAL       421    421       IRON (HEME B AXIAL LIGAND) (PROBABLE).
FT   CROSSLNK    284    288       1'-histidyl-3'-tyrosine (His-Tyr)
FT                                (By similarity).
SQ   SEQUENCE   659 AA;  75028 MW;  23D6FB4B04732D23 CRC64;

  Query Match             63.2%;  Score 36;  DB 1;  Length 659;
  Best Local Similarity   75.0%;  Pred. No. 28;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIH 8
              ||||::||
Db        576 KKQYSAIH 583


RESULT 15
YRD6_CAEEL
ID   YRD6_CAEEL     STANDARD;      PRT;  1268 AA.
AC   Q09575;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Hypothetical protein K02A2.6 in chromosome II.
GN   K02A2.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Hallsworth K.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: WEAK, TO RETROVIRUS-RELATED POL POLYPROTEIN.
CC   -!- SIMILARITY: Contains 1 CCHC-type zinc finger.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U23171; AAC46707.1; -.
DR   PIR; B88209; B88209.
DR   WormPep; K02A2.6; CE02792.
DR   InterPro; IPR009007; Pept_A_acid.
DR   InterPro; IPR001584; Rve.
DR   InterPro; IPR000477; RVTse.
DR   InterPro; IPR001878; Znf_CCHC.
DR   Pfam; PF00665; rve; 1.
DR   Pfam; PF00078; rvt; 1.
DR   Pfam; PF00098; zf-CCHC; 1.
DR   PRINTS; PR00939; C2HCZNFINGER.
DR   SMART; SM00343; ZnF_C2HC; 1.
DR   PROSITE; PS50158; ZF_CCHC; 1.
KW   Hypothetical protein; Zinc-finger.
FT   ZN_FING     239    256       CCHC-TYPE.
SQ   SEQUENCE   1268 AA;  143493 MW;  AACCAC4235AAE468 CRC64;

  Query Match             63.2%;  Score 36;  DB 1;  Length 1268;
  Best Local Similarity   50.0%;  Pred. No. 56;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              :| |  :|||
Db       1043 QKVYVQVHHG 1052


Search completed: October  4, 2004, 18:57:07
Job time : 10.2609 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:37:37 ; Search time 47.3913 Seconds
(without alignments)
66.577 Million cell updates/sec

Title: US-10-046-209-3
Perfect score: 57
Sequence: 1 KKQYTSIHHG 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	57	100.0	49	6	O97917	O97917 bos taurus
2	57	100.0	113	13	Q8JH58	Q8jh58 chelydra se
3	57	100.0	218	11	Q8BPV5	Q8bpv5 mus musculu
4	57	100.0	239	13	Q8UUI7	Q8uui7 brachydanio
5	57	100.0	357	13	Q8UUI8	Q8uui8 brachydanio
6	57	100.0	384	11	Q8BPC7	Q8bpc7 mus musculu
7	57	100.0	472	13	Q8UUS0	Q8uus0 brachydanio
8	57	100.0	534	13	O93296	O93296 gallus gall
9	57	100.0	569	13	Q9PVL1	Q9pvl1 gallus gall
10	57	100.0	612	13	Q9I9E7	Q9i9e7 brachydanio
11	57	100.0	678	13	Q7ZZT1	Q7zzt1 brachydanio
12	57	100.0	694	13	Q8UUR9	Q8uur9 brachydanio
13	57	100.0	695	13	Q9DGJ8	Q9dgj8 gallus gall
14	57	100.0	699	13	O57394	O57394 narke japon
15	57	100.0	738	13	Q90W28	Q90w28 brachydanio
16	57	100.0	751	13	Q9DGJ7	Q9dgj7 gallus gall
17	54	94.7	693	13	Q98SG0	Q98sg0 xenopus lae
18	54	94.7	695	13	Q98SF9	Q98sf9 xenopus lae
19	54	94.7	695	13	Q7ZXQ0	Q7zxq0 xenopus lae
20	54	94.7	747	13	Q91963	Q91963 xenopus. ap
21	50	87.7	82	4	Q16020	Q16020 homo sapien
22	50	87.7	82	4	Q16014	Q16014 homo sapien
23	50	87.7	82	4	Q16019	Q16019 homo sapien
24	42	73.7	133	5	O17443	O17443 brevicoryne
25	38	66.7	228	16	Q8A0P0	Q8a0p0 bacteroides
26	38	66.7	289	16	Q8RIL2	Q8ril2 fusobacteri
27	37	64.9	164	16	Q81P90	Q81p90 bacillus an
28	37	64.9	442	16	Q92AT6	Q92at6 listeria in
29	37	64.9	2584	5	Q9TW88	Q9tw88 caenorhabdi
30	37	64.9	2606	5	Q21920	Q21920 caenorhabdi
31	36	63.2	71	16	Q83DX8	Q83dx8 coxiella bu
32	36	63.2	116	11	Q8CDE5	Q8cde5 mus musculu
33	36	63.2	159	17	Q8TTR5	Q8ttr5 methanosarc
34	36	63.2	159	17	Q8TTR3	Q8ttr3 methanosarc
35	36	63.2	275	16	Q9RXF2	Q9rxf2 deinococcus
36	36	63.2	333	13	Q90YZ9	Q90yz9 sphyraena i
37	36	63.2	412	5	Q21806	Q21806 caenorhabdi
38	36	63.2	443	2	Q9ZA32	Q9za32 streptomyce
39	36	63.2	464	10	Q9FEP0	Q9fep0 adonis pala
40	36	63.2	550	8	Q7YAJ6	Q7yaj6 chara vulga
41	36	63.2	729	16	Q8DKG0	Q8dkg0 synechococc
42	36	63.2	904	10	Q9SF47	Q9sf47 arabidopsis
43	36	63.2	944	5	Q8T4F6	Q8t4f6 drosophila
44	36	63.2	944	5	Q9VKQ6	Q9vkq6 drosophila
45	36	63.2	1313	5	P91422	P91422 caenorhabdi

ALIGNMENTS

RESULT 1
O97917
ID O97917 PRELIMINARY; PRT; 49 AA.
AC O97917;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```
DE   Amyloid protein (Fragment).
GN   APP.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20063685; PubMed=10594237;
RA   Konfortov B.A., Licence V.E., Miller J.R.;
RT   "Resequencing of DNA from a diverse panel of cattle reveals a high
RT   level of polymorphism in both intron and exon.";
RL   Mamm. Genome 10:1142-1145(1999).
DR   EMBL; AJ133033; CAB38017.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
FT   NON_TER       1      1
FT   NON_TER      49     49
SQ   SEQUENCE   49 AA;  5183 MW;  6287463F0559BDED CRC64;

  Query Match             100.0%;  Score 57;  DB 6;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 0.0018;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db         37 KKQYTSIHHG 46


RESULT 2
Q8JH58
ID   Q8JH58      PRELIMINARY;      PRT;   113 AA.
AC   Q8JH58;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid beta protein (Fragment).
OS   Chelydra serpentina serpentina (common snapping turtle).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Testudines; Cryptodira; Testudinoidea; Chelydridae; Chelydra.
OX   NCBI_TaxID=134619;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21876906; PubMed=11882478;
RA   Trudeau V.L., Chiu S., Kennedy S.W., Brooks R.J.;
RT   "Octylphenol (OP) alters the expression of members of the amyloid
RT   protein family in the hypothalamus of the snapping turtle, Chelydra
RT   serpentina serpentina.";
RL   Environ. Health Perspect. 110:269-275(2002).
DR   EMBL; AF541917; AAN04908.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   113 AA;  12750 MW;  72515C930496E053 CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 113;
  Best Local Similarity   100.0%;  Pred. No. 0.0042;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db         68 KKQYTSIHHG 77


RESULT 3
Q8BPV5
ID   Q8BPV5      PRELIMINARY;      PRT;   218 AA.
AC   Q8BPV5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid beta (Fragment).
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK052448; BAC34997.1; -.
DR   MGD; MGI:88059; App.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   218 AA;  24118 MW;  95B55AFDAE1D0EF5 CRC64;

  Query Match             100.0%;  Score 57;  DB 11;  Length 218;
  Best Local Similarity   100.0%;  Pred. No. 0.0082;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        173 KKQYTSIHHG 182


RESULT 4
```

```
Q8UUI7
ID   Q8UUI7      PRELIMINARY;      PRT;   239 AA.
AC   Q8UUI7;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative membrane protein (Fragment).
GN   APPB.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315638; CAC85735.1; -.
DR   ZFIN; ZDB-GENE-020220-1; appb.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   239 AA;  27048 MW;  8A69F746F821BAE2 CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 239;
  Best Local Similarity   100.0%;  Pred. No. 0.009;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        194 KKQYTSIHHG 203


RESULT 5
Q8UUI8
ID   Q8UUI8      PRELIMINARY;      PRT;   357 AA.
AC   Q8UUI8;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative mebrane protein (Fragment).
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Embryo;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315637; CAC85734.1; -.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   357 AA;  40962 MW;  07D99EEF6C55B2D8 CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 357;
  Best Local Similarity   100.0%;  Pred. No. 0.014;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        312 KKQYTSIHHG 321


RESULT 6
Q8BPC7
ID   Q8BPC7      PRELIMINARY;      PRT;   384 AA.
AC   Q8BPC7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid beta (Fragment).
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK076506; BAC36369.1; -.
DR   MGD; MGI:88059; App.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
```

```
SQ   SEQUENCE   384 AA;  43990 MW;  A81B1AD8AE683173 CRC64;

  Query Match             100.0%;  Score 57;  DB 11;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        339 KKQYTSIHHG 348


RESULT 7
Q8UUS0
ID   Q8UUS0      PRELIMINARY;      PRT;   472 AA.
AC   Q8UUS0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative membrane protein (Fragment).
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315636; CAC85733.1; -.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   472 AA;  53787 MW;  24F7128BE3356550 CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 472;
  Best Local Similarity   100.0%;  Pred. No. 0.018;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        427 KKQYTSIHHG 436


RESULT 8
O93296
ID   O93296      PRELIMINARY;      PRT;   534 AA.
AC   O93296;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98337885; PubMed=9671674;
RA   Barnes N.Y., Li L., Yoshikawa K., Schwartz L.M., Oppenheim R.W.,
RA   Milligan C.E.;
RT   "Increased production of amyloid precursor protein provides a
RT   substrate for caspase-3 in dying motoneurons.";
RL   J. Neurosci. 18:5869-5880(1998).
DR   EMBL; AF042098; AAC25052.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   534 AA;  60597 MW;  FB53ECC2E66D4C92 CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        489 KKQYTSIHHG 498


RESULT 9
Q9PVL1
ID   Q9PVL1      PRELIMINARY;      PRT;   569 AA.
AC   Q9PVL1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
GN   APP.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
```

```
RA   Coulson E.J., Paliga K., Beyreuther K., Masters C.L.;
RT   "What the evolution of the amyloid protein precursor supergene family
RT   tells us about its function.";
RL   Neurochem. Int. 0:0-0(2000).
DR   EMBL; AF030341; AAF12698.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   569 AA;  64753 MW;  0AB8BB851863A19D CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 569;
  Best Local Similarity   100.0%;  Pred. No. 0.022;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        525 KKQYTSIHHG 534


RESULT 10
Q9I9E7
ID   Q9I9E7      PRELIMINARY;      PRT;   612 AA.
AC   Q9I9E7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Slavov D.B., Gardiner K.;
RT   "An App cDNA from Zebrafish (Danio rerio).";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF257742; AAF71748.1; -.
DR   HSSP; P05067; 1HZ3.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   612 AA;  69710 MW;  59A9ACBDF9C59EFF CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 612;
  Best Local Similarity   100.0%;  Pred. No. 0.024;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        567 KKQYTSIHHG 576


RESULT 11
Q7ZZT1
ID   Q7ZZT1      PRELIMINARY;      PRT;   678 AA.
AC   Q7ZZT1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid protein a variant 2.
GN   APPA.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Groth C., Lardelli M.;
RT   "Investigation of zebrafish appa expression during embryogenesis.";
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY271746; AAP22958.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   678 AA;  76755 MW;  94163778444FD0BC CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 678;
  Best Local Similarity   100.0%;  Pred. No. 0.026;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        633 KKQYTSIHHG 642


RESULT 12
Q8UUR9
ID   Q8UUR9      PRELIMINARY;      PRT;   694 AA.
```

```
AC   Q8UUR9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative membrane protein.
GN   APPB.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   PubMed=11862463;
RA   Musa A., Lehrach H., Russo V.E.A.;
RT   "Distinct expression patterns of two zebrafish homologues of the human
RT   APP gene during embryonic development.";
RL   Dev. Genes Evol. 211:563-567(2001).
DR   EMBL; AJ315639; CAC85736.1; -.
DR   ZFIN; ZDB-GENE-020220-1; appb.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   694 AA;  79228 MW;  2B03382D411162DC CRC64;

  Query Match           100.0%;  Score 57;  DB 13;  Length 694;
  Best Local Similarity  100.0%;  Pred. No. 0.027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        649 KKQYTSIHHG 658


RESULT 13
Q9DGJ8
ID   Q9DGJ8      PRELIMINARY;      PRT;   695 AA.
AC   Q9DGJ8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Beta-amyloid precursor protein 695 isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sarasa M., Rodolosse A., Sorribas V.;
RT   "Cloning of full-length chicken beta-amyloid precursor protein
RT   isoforms.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289218; AAG00593.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   695 AA;  78565 MW;  F201ED02AEC86D95 CRC64;

  Query Match           100.0%;  Score 57;  DB 13;  Length 695;
  Best Local Similarity  100.0%;  Pred. No. 0.027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        650 KKQYTSIHHG 659


RESULT 14
O57394
ID   O57394      PRELIMINARY;      PRT;   699 AA.
AC   O57394;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   EL amyloid precursor protein 699.
GN   EL APP699.
OS   Narke japonica (Electric ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Torpediniformes; Narcinoidei; Narkidae; Narke.
OX   NCBI_TaxID=62965;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Electric lobe;
RX   MEDLINE=98129705; PubMed=9461486;
RA   Iijima K., Lee D.-S., Okutsu J., Tomita S., Hirashima N., Kirino Y.,
RA   Suzuki T.;
RT   "cDNA isolation of Alzheimer's amyloid precursor protein from
RT   cholinergic nerve terminals of the electric organ of the electric
RT   ray.";
RL   Biochem. J. 330:29-33(1998).
DR   EMBL; AB005544; BAA24230.1; -.
DR   HSSP; P05067; 1HZ3.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
```

```
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   699 AA;  78879 MW;  952915C309D50E5C CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 699;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        654 KKQYTSIHHG 663


RESULT 15
Q90W28
ID   Q90W28      PRELIMINARY;      PRT;   738 AA.
AC   Q90W28;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid precursor protein.
GN   APPA OR APP.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Groth C., Lardelli M.;
RT   "Expression analysis of zebrafish app.";
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF389401; AAK64495.1; -.
DR   ZFIN; ZDB-GENE-000616-13; appa.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Protease inhibitor; Serine protease inhibitor.

SQ   SEQUENCE   738 AA;  83577 MW;  AF480F6D308FD298 CRC64;

  Query Match             100.0%;  Score 57;  DB 13;  Length 738;
  Best Local Similarity   100.0%;  Pred. No. 0.029;
  Matches   10;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKQYTSIHHG 10
              ||||||||||
Db        693 KKQYTSIHHG 702


Search completed: October  4, 2004, 18:56:22
Job time : 48.3913 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:37:13 ; Search time 63.8478 Seconds
(without alignments)
48.679 Million cell updates/sec

Title: US-10-048-209-4
Perfect score: 62
Sequence: 1 KKKQYTSIHHG 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	13	4	AAB67794	Aab67794 Cytoplasm
2	62	100.0	15	4	AAB67798	Aab67798 Cytoplasm
3	62	100.0	19	7	ADC49255	Adc49255 Human inh
4	62	100.0	32	6	AAO19883	Aao19883 Human amy
5	62	100.0	33	7	ADC49254	Adc49254 Human inh
6	62	100.0	41	4	AAM16658	Aam16658 Peptide #
7	62	100.0	41	4	ABB35642	Abb35642 Peptide #
8	62	100.0	41	4	AAM29142	Aam29142 Peptide #
9	62	100.0	41	4	ABB30475	Abb30475 Peptide #
10	62	100.0	41	4	ABB21071	Abb21071 Protein #
11	62	100.0	41	4	AAM56458	Aam56458 Human bra
12	62	100.0	41	4	AAM04374	Aam04374 Peptide #
13	62	100.0	41	5	ABG38416	Abg38416 Human pep
14	62	100.0	44	2	AAW53985	Aaw53985 Human ALZ
15	62	100.0	47	2	AAR58917	Aar58917 Cytoplasm
16	62	100.0	47	2	AAW26395	Aaw26395 Amyloid p
17	62	100.0	47	2	AAW26402	Aaw26402 Amyloid p
18	62	100.0	47	2	AAW26400	Aaw26400 Amyloid p
19	62	100.0	47	2	AAW26403	Aaw26403 Amyloid p
20	62	100.0	47	2	AAW26399	Aaw26399 Amyloid p
21	62	100.0	47	2	AAW26401	Aaw26401 Amyloid p
22	62	100.0	47	2	AAW26520	Aaw26520 Amyloid p
23	62	100.0	47	2	AAW26518	Aaw26518 Amyloid p
24	62	100.0	47	2	AAW26521	Aaw26521 Amyloid p
25	62	100.0	47	2	AAW26519	Aaw26519 Amyloid p
26	62	100.0	47	2	AAW26513	Aaw26513 Amyloid p
27	62	100.0	47	2	AAW26517	Aaw26517 Amyloid p
28	62	100.0	47	2	AAW42984	Aaw42984 APP isofo
29	62	100.0	47	2	AAW42986	Aaw42986 APP isofo
30	62	100.0	47	2	AAW42987	Aaw42987 APP isofo
31	62	100.0	47	2	AAW42983	Aaw42983 APP isofo
32	62	100.0	47	2	AAW42985	Aaw42985 APP isofo
33	62	100.0	47	2	AAW44755	Aaw44755 APP-REP 7
34	62	100.0	47	2	AAW44753	Aaw44753 APP-REP 7
35	62	100.0	47	2	AAW44749	Aaw44749 APP-REP 7
36	62	100.0	47	2	AAW44756	Aaw44756 APP-REP 7
37	62	100.0	47	2	AAW44754	Aaw44754 APP-REP 7
38	62	100.0	47	2	AAW44757	Aaw44757 APP-REP 7
39	62	100.0	47	2	AAW44274	Aaw44274 Amyloid p
40	62	100.0	47	4	AAB67791	Aab67791 Cytoplasm
41	62	100.0	47	7	ADC49250	Adc49250 Human inh
42	62	100.0	47	7	ADC49249	Adc49249 Inhibitor
43	62	100.0	47	7	ADC49251	Adc49251 Human inh
44	62	100.0	49	2	AAR35087	Aar35087 Human amy
45	62	100.0	49	4	AAM14458	Aam14458 Peptide #

ALIGNMENTS

RESULT 1
AAB67794
ID AAB67794 standard; peptide; 13 AA.
XX
AC AAB67794;
XX
DT 11-JUN-2001 (first entry)
XX
DE Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1

```
FT                   /note= "this residue represents an internalisation
FT                   peptide such as the sequence given in AAB67795"
FT   Misc-difference 13
FT                   /note= "this residue represents V, VV, VVE, VVEV, VVEVD"
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Claim 1; Page 13; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 13 AA;

  Query Match             100.0%;  Score 62;  DB 4;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 0.00019;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db          2 KKKQYTSIHHG 12


RESULT 2
AAB67798
ID   AAB67798 standard; peptide; 15 AA.
XX
AC   AAB67798;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Cytoplasmic domain of the amyloid protein precursor (APP).
XX
KW   Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO200109170-A1.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-FR002174.
XX
PR   30-JUL-1999;   99FR-00009929.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
XX
PI   Allinquant B,  Prochiantz A;
XX
DR   WPI; 2001-257398/26.
XX
PT   Peptides derived from the cytoplasmic domain of the amyloid protein
PT   precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS   Disclosure; Page 2; 28pp; French.
XX
CC   The present sequence represents a peptide derived from the cytoplasmic
CC   domain of the human amyloid protein precursor (APP). APP peptides derived
CC   from the cytoplasmic domain, and containing the membrane domain
CC   juxtaposed to the cytoplasmic domain of APP are useful for selecting and
CC   screening products capable of inhibiting apoptosis. The peptides are
CC   useful in the treatment of cancer and Alzheimer's disease
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 62;  DB 4;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00022;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db          1 KKKQYTSIHHG 11


RESULT 3
ADC49255
ID   ADC49255 standard; protein; 19 AA.
XX
AC   ADC49255;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human inhibitor of metabolic degradation of APP variant #4.
XX
KW   human; inhibitory factor; metabolic degradation of APP;
KW   amyloid precursor protein; Alzheimer's disease; mutant; mutein.
XX
OS   Homo sapiens.
XX
PN   JP2002360252-A.
XX
PD   17-DEC-2002.
XX
PF   27-APR-2001; 2001JP-00133178.
XX
```

PR 27-APR-2001; 2001JP-00133178.
XX
PA (SUZU/) SUZUKI T.
PA (SUMU) SUMITOMO SEIYAKU KK.
XX
DR WPI; 2003-516151/49.
XX
PT An inhibitory factor of metabolic degradation of amyloid precursor
PT protein (APP) which inhibits formation of beta-amyloid, useful in the
PT treatment of Alzheimer's disease.
XX
PS Disclosure; SEQ ID NO 15; 33pp; Japanese.
XX
CC The invention relates to an inhibitory factor of metabolic degradation of
CC APP. The factor is useful in the treatment of Alzheimer's disease. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 19 AA;

  Query Match   100.0%; Score 62; DB 7; Length 19;
  Best Local Similarity   100.0%; Pred. No. 0.00029;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKKQYTSIHHG 11
              |||||||||||
Db   1 KKKQYTSIHHG 11

RESULT 4
AAO19883
ID AAO19883 standard; peptide; 32 AA.
XX
AC AAO19883;
XX
DT 11-AUG-2003 (first entry)
XX
DE Human amyloid precursor protein APP immunogenic peptide #3.
XX
KW Human; APP; amyloid precursor protein; immunogen; Alzheimer's disease;
KW high-throughput screening; neuroprotective; nootropic; antiparkinsonian.
XX
OS Homo sapiens.
XX
PN WO2003001881-A2.
XX
PD 09-JAN-2003.
XX
PF 26-JUN-2002; 2002WO-US020267.
XX
PR 26-JUN-2001; 2001US-0300959P.
XX
PA (NYME-) NEW YORK STATE OFFICE MENTAL HEALTH.
XX
PI Mathews PM, Nixon RA, Schmidt SD, Jiang Y;
XX
DR WPI; 2003-210182/20.
XX
PT Identifying compounds that modulates the generation of metabolites
PT associated with a disease or disorder, for treating e.g. Alzheimer's
PT disease by determining levels of a cellular component protein, or its
PT conformation state.
XX
PS Example 1; Page 29; 69pp; English.
XX
CC The present invention relates to a method of identifying compounds that
CC modulate the generation of one or more metabolites associated with a
CC disease or disorder comprising determining levels of a cellular component
CC protein or a conformation state of a cellular precursor protein. In
CC particular, the method can be used to determine levels of amyloid
CC precursor protein (APP), which is associated with Alzheimer's disease. It
CC is also useful for identifying compounds as drugs for treating diseases
CC or disorders associated with metabolic and/or proteolytic pathways, e.g.
CC Alzheimer's disease, Parkinson's disease, Huntington's disease, lysosomal
CC storage disorders, prion diseases, the tau-based neurodegenerative
CC disorders, and other non-AD amyloidoses. The present sequence is an
CC immunogenic portion of human APP
XX
SQ Sequence 32 AA;

  Query Match   100.0%; Score 62; DB 6; Length 32;
  Best Local Similarity   100.0%; Pred. No. 0.00051;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKKQYTSIHHG 11
              |||||||||||
Db   5 KKKQYTSIHHG 15

RESULT 5
ADC49254
ID ADC49254 standard; protein; 33 AA.
XX
AC ADC49254;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human inhibitor of metabolic degradation of APP variant #3.
XX
KW human; inhibitory factor; metabolic degradation of APP;
KW amyloid precursor protein; Alzheimer's disease; mutant; mutein.
XX
OS Homo sapiens.
XX
PN JP2002360252-A.
XX
PD 17-DEC-2002.
XX
PF 27-APR-2001; 2001JP-00133178.
XX
PR 27-APR-2001; 2001JP-00133178.
XX
PA (SUZU/) SUZUKI T.
PA (SUMU) SUMITOMO SEIYAKU KK.
XX

DR WPI; 2003-516151/49.
XX
PT An inhibitory factor of metabolic degradation of amyloid precursor
PT protein (APP) which inhibits formation of beta-amyloid, useful in the
PT treatment of Alzheimer's disease.
XX
PS Disclosure; SEQ ID NO 14; 33pp; Japanese.
XX
CC The invention relates to an inhibitory factor of metabolic degradation of
CC APP. The factor is useful in the treatment of Alzheimer's disease. The
CC present sequence is used in the exemplification of the invention.
XX
SQ Sequence 33 AA;

Query Match 100.0%; Score 62; DB 7; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.00053;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKQYTSIHHG 11
||||||||||||
Db 1 KKKQYTSIHHG 11

RESULT 6
AAM16658
ID AAM16658 standard; protein; 41 AA.
XX
AC AAM16658;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #3092 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 27; SEQ ID NO 21484; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

Query Match 100.0%; Score 62; DB 4; Length 41;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKQYTSIHHG 11
||||||||||||
Db 28 KKKQYTSIHHG 38

RESULT 7
ABB35642
ID ABB35642 standard; peptide; 41 AA.
XX
AC ABB35642;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #3148 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 27; SEQ ID NO 28277; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 62; DB 4; Length 41;
  Best Local Similarity 100.0%; Pred. No. 0.00067;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKKQYTSIHHG 11
              ||||||||||
Db   28 KKKQYTSIHHG 38

RESULT 8
AAM29142
ID AAM29142 standard; protein; 41 AA.
XX
AC AAM29142;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #3179 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 27; SEQ ID NO 29411; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 62; DB 4; Length 41;
  Best Local Similarity 100.0%; Pred. No. 0.00067;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKKQYTSIHHG 11
              ||||||||||
Db   28 KKKQYTSIHHG 38

RESULT 9
ABB30475
ID ABB30475 standard; peptide; 41 AA.
XX
AC ABB30475;
XX
DT 01-FEB-2002 (first entry)
XX
DE Peptide #3126 encoded by breast cell single exon nucleic acid probe.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
KW cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
PS Claim 27; SEQ ID NO 13443; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 62; DB 4; Length 41;
  Best Local Similarity   100.0%; Pred. No. 0.00067;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKKQYTSIHHG 11
              |||||||||||
Db   28 KKKQYTSIHHG 38

RESULT 10
ABB21071
ID ABB21071 standard; protein; 41 AA.
XX
AC ABB21071;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #3070 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 22841; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 62; DB 4; Length 41;
  Best Local Similarity   100.0%; Pred. No. 0.00067;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKKQYTSIHHG 11
              |||||||||||
Db   28 KKKQYTSIHHG 38

RESULT 11
AAM56458
ID AAM56458 standard; protein; 41 AA.
XX

AC AAM56458;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 28563.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 28563; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 62; DB 4; Length 41;
  Best Local Similarity 100.0%; Pred. No. 0.00067;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KKKQYTSIHHG 11
              |||||||||||
Db   28 KKKQYTSIHHG 38

RESULT 12
AAM04374

ID AAM04374 standard; protein; 41 AA.
XX
AC AAM04374;
XX
DT 09-OCT-2001 (first entry)
XX
DE Peptide #3056 encoded by probe for measuring breast gene expression.
XX
KW Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 27; SEQ ID NO 13114; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 41 AA;

  Query Match   100.0%; Score 62; DB 4; Length 41;
  Best Local Similarity 100.0%; Pred. No. 0.00067;
  Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 KKKQYTSIHHG 11
              ||||||||||
Db         28 KKKQYTSIHHG 38
```

```
RESULT 13
ABG38416
ID   ABG38416 standard; peptide; 41 AA.
XX
AC   ABG38416;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human peptide encoded by genome-derived single exon probe SEQ ID 28081.
XX
KW   Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US000665.
XX
PR   04-FEB-2000; 2000US-0180312P.
PR   26-MAY-2000; 2000US-0207456P.
PR   30-JUN-2000; 2000US-00608408.
PR   03-AUG-2000; 2000US-00632366.
PR   21-SEP-2000; 2000US-0234687P.
PR   27-SEP-2000; 2000US-0236359P.
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples.
XX
PS   Claim 27; SEQ ID NO 28081; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of probes
CC   ; the novel set of probes which hybridise at high stringency to a nucleic
CC   acid expressed in the human lung; measuring gene expression in a sample
CC   derived from human lung, comprising (a) contacting the array with a
CC   collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of the
CC   array; identifying exons in a eukaryotic genome, comprising (a)
CC   algorithmically predicting at least one exon from genomic sequences of
CC   the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene expression
CC   analysis, and for identifying exons in a gene, particularly using human
CC   lung derived mRNA and for the study of lung diseases such as asthma, lung
CC   cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC   disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC   tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC   Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC   histiocytosis, lymphangioleiomyomtosis, pulmonary alveolar proteinosis,
CC   Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC   dyskinesis, pulmonary hypertension and hyaline membrane disease. The
CC   present sequence is a peptide/protein encoded by a single exon probe of
CC   the invention. Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 41 AA;
```

```
  Query Match             100.0%;  Score 62;  DB 5;  Length 41;
  Best Local Similarity   100.0%;  Pred. No. 0.00067;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||
Db         28 KKKQYTSIHHG 38
```

```
RESULT 14
AAW53985
ID   AAW53985 standard; protein; 44 AA.
XX
AC   AAW53985;
XX
DT   18-AUG-1998  (first entry)
XX
DE   Human ALZASp2.
XX
KW   Dsas; DSASp; alzas; Down's syndrome; diagnosis; therapy; human;
```

```
KW   Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO9807850-A2.
XX
PD   26-FEB-1998.
XX
PF   22-AUG-1997;   97WO-EP004599.
XX
PR   22-AUG-1996;   96CA-02183901.
XX
PA   (BERG/) BERGMANN J E.
PA   (PRED/) PREDDIE E R.
XX
PI   Bergmann JE,  Preddie ER;
XX
DR   WPI; 1998-169155/15.
DR   N-PSDB; AAV23755.
XX
PT   Nucleic acid molecules dsas, and alzas - used for detecting and treating
PT   Down's syndrome and Alzheimer's disease.
XX
PS   Claim 13; Fig 1M; 96pp; English.
XX
CC   This sequence is the ALZASp2 encoded by the nucleic acid alzas. The dsas
CC   and alzas DNA sequences are the nucleic acids of the invention. Reagents
CC   specifically for DSASp can be used for the diagnosis of Down's syndrome
CC   in humans and especially in pregnant women. Molecules that inhibit the
CC   activity of the promoters (PDS1, PDS2, PDS3, and PDS4) for dsas can be
CC   used for treating Down's syndrome. The reagent capable of detecting alzas
CC   can be used for detecting Alzheimer's disease, especially in the pre-
CC   symptomatic stage. Substances that inhibit the promoters for alzas can be
CC   used in treating Alzheimer's disease
XX
SQ   Sequence 44 AA;

  Query Match             100.0%;  Score 62;  DB 2;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 0.00072;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db         19 KKKQYTSIHHG 29


RESULT 15
AAR58917
ID   AAR58917 standard; peptide; 47 AA.
XX
AC   AAR58917;
XX
DT   25-MAR-2003  (revised)
DT   15-APR-1995  (first entry)
XX
DE   Cytoplasmic tail portion of APP695 (residues 649-695).
XX
KW   Amyloid precursor protein; APP695; beta amyloid; cytoplasmic tail;
KW   Alzheimer's disease.
XX
OS   Homo sapiens.
XX
PN   WO9419692-A1.
XX
PD   01-SEP-1994.
XX
PF   17-FEB-1994;   94WO-US001712.
XX
PR   18-FEB-1993;   93US-00019208.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
XX
PI   Nishimoto I;
XX
DR   WPI; 1994-294486/36.
XX
PT   Identifying cpds. useful for treating or preventing Alzheimer's disease -
PT   by determining whether it interferes with the association of the couplone
PT   portion of amyloid precursor protein to G polypeptide.
XX
PS   Claim 4; Page 29; 71pp; English.
XX
CC   Beta amyloid is synthesised as part of a larger protein referred to as
CC   amyloid precursor protein (APP), which has a number of isoforms in
CC   humans, including APP695 and APP770. The amino terminal of beta amyloid
CC   is generated by cleavage of a peptide bond of APP which in APP695 lies
CC   between Met596 and Asp597. APP forms a complex with Go, a GTP-binding
CC   protein (or "G protein") in brain. Go is made of one alpha subunit and
CC   one Beta-gamma subunit. Two isoforms of Go, known as Go1 (or GoA) and Go2
CC   (or GoB) have been identified; they have slight AA differences in their
CC   alpha subunits. The cDNA sequence and deduced AA sequence of the alpha
CC   subunits in each of Go1 and Go2 are shown in AAQ69002/R58914 and
CC   AAQ69004/R58924 respectively. The cytoplasmic App695 sequence His657-
CC   Lys676 (AAR58913) possesses a specific Go-activating function, and is
CC   necessary for complex formation of this APP with Go. The cytoplasmic tail
CC   portion of APP695 from residues 649-695 and the membrane-spainning
CC   portion of APP695 from residues 639-648, and the entire membrane-spanning
CC   segment from residues 625-648 are given in AAR58917, AAR58918, and
CC   AAR58922 respectively. The sequence of all of APP is given in
CC   AAQ69003/R58923. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 47 AA;

  Query Match             100.0%;  Score 62;  DB 2;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 0.00078;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db          1 KKKQYTSIHHG 11


Search completed: October  4, 2004, 18:52:37
Job time : 64.8478 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:45:55 ; Search time 16.2609 Seconds
(without alignments)
65.071 Million cell updates/sec

Title: US-10-048-209-4
Perfect score: 62
Sequence: 1 KKKQYTSIHHG 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	82	2	PQ0438	Alzheimer's diseas
2	62	100.0	695	1	A49795	Alzheimer's diseas
3	62	100.0	695	2	A27485	Alzheimer's diseas
4	62	100.0	695	2	S00550	Alzheimer's diseas
5	62	100.0	770	1	QRHUA4	Alzheimer's diseas
6	59	95.2	747	2	JH0773	Alzheimer's diseas
7	43	69.4	406	2	S76307	hypothetical prote
8	40	64.5	452	2	T04781	hypothetical prote
9	39	62.9	452	2	S58994	NADH2 dehydrogenas
10	38	61.3	488	2	S64140	hypothetical prote
11	37	59.7	142	2	T10871	y4kQ protein - Rhi
12	37	59.7	191	2	A35981	sperm membrane pro
13	37	59.7	267	2	T52381	zinc finger protei
14	37	59.7	391	2	H96517	protein T2E6.22 [i
15	37	59.7	442	2	AH1661	ATP-dependent RNA
16	37	59.7	511	2	JC1404	CDEI-box DNA-bindi
17	37	59.7	653	2	A46362	amyloid precursor-
18	37	59.7	751	2	A49974	beta-amyloid precu
19	37	59.7	763	2	A49321	amyloid beta (A4)
20	37	59.7	765	2	S42880	amyloid precursor-
21	37	59.7	2584	2	T24158	hypothetical prote
22	37	59.7	2606	2	T24157	hypothetical prote
23	36	58.1	119	2	AG3136	hypothetical prote
24	36	58.1	275	2	C75529	competence protein
25	36	58.1	412	2	T24023	hypothetical prote
26	36	58.1	443	2	T46532	probable dTDP-4-ke
27	36	58.1	693	2	T49296	hypothetical prote
28	36	58.1	749	2	B69941	ATP-dependent RNA
29	36	58.1	1268	2	B88209	protein K02A2.6 [i
30	36	58.1	1313	2	T29193	hypothetical prote
31	36	58.1	3973	2	B71612	hypothetical prote
32	35	56.5	102	2	T07078	cold stress protei
33	35	56.5	150	2	F96924	flavodoxin [import
34	35	56.5	177	2	F84245	inosine-5'-monopho
35	35	56.5	211	2	T01627	probable ATP bindi
36	35	56.5	239	2	T01463	hypothetical prote
37	35	56.5	264	2	T15289	hypothetical prote
38	35	56.5	266	2	I49114	Ly49H - mouse
39	35	56.5	269	2	T16910	hypothetical prote
40	35	56.5	332	2	T25779	hypothetical prote
41	35	56.5	361	2	D83798	phosphoserine amin
42	35	56.5	410	2	AF1660	aminopeptidases ho
43	35	56.5	410	2	AG1288	aminopeptidases ho
44	35	56.5	412	2	T23385	hypothetical prote
45	35	56.5	443	2	B39794	transcription fact

ALIGNMENTS

RESULT 1
PQ0438
Alzheimer's disease amyloid A4 protein precursor - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-1993 #sequence_revision 19-Oct-1995 #text_change 19-Oct-1995
C;Accession: PQ0438; C60045
R;Davidson, J.S.; West, R.L.; Kotikalapudi, P.; Maroun, L.E.
Biochem. Biophys. Res. Commun. 188, 905-911, 1992
A;Title: Sequence and methylation in the beta/A4 region of the rabbit amyloid precursor protein gene.
A;Reference number: PQ0438; MUID:93075180; PMID:1445331
A;Accession: PQ0438
A;Molecule type: DNA
A;Residues: 1-82 <DAV>
A;Cross-references: GB:M83558; GB:M83657
R;Johnstone, E.M.; Chaney, M.O.; Norris, F.H.; Pascual, R.; Little, S.P.
Brain Res. Mol. Brain Res. 10, 299-305, 1991
A;Title: Conservation of the sequence of the Alzheimer's disease amyloid peptide in dog, polar bear and five other mammals by cross-species polymerase chain reaction analysis.

A;Reference number: A60045; MUID:92017079; PMID:1656157
A;Accession: C60045
A;Molecule type: mRNA
A;Residues: 12-68 <JOH>
A;Cross-references: EMBL:X56129
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; Alzheimer's disease; amyloid; Down's syndrome

  Query Match 100.0%; Score 62; DB 2; Length 82;
  Best Local Similarity 100.0%; Pred. No. 0.0002;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKQYTSIHHG 11
      |||||||||||
Db 69 KKKQYTSIHHG 79

RESULT 2
A49795
Alzheimer's disease amyloid beta protein precursor - crab-eating macaque
C;Species: Macaca fascicularis (crab-eating macaque)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A49795
R;Podlisny, M.B.; Tolan, D.R.; Selkoe, D.J.
Am. J. Pathol. 138, 1423-1435, 1991
A;Title: Homology of the amyloid beta protein precursor in monkey and human supports a primate model for beta amyloidosis in Alzheimer's disease.
A;Reference number: A49795; MUID:91273117; PMID:1905108
A;Accession: A49795
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-695 <POD>
A;Cross-references: GB:M58727; NID:g342062; PIDN:AAA36829.1; PID:g342063
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing

  Query Match 100.0%; Score 62; DB 1; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.0019;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKQYTSIHHG 11
      |||||||||||
Db 649 KKKQYTSIHHG 659

RESULT 3
A27485
Alzheimer's disease amyloid beta/A4 protein homolog precursor - mouse
N;Alternate names: proteinase nexin II
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 13-Aug-1999
C;Accession: A27485; S19727; I49485
R;Yamada, T.; Sasaki, H.; Furuya, H.; Miyata, T.; Goto, I.; Sakaki, Y.
Biochem. Biophys. Res. Commun. 149, 665-671, 1987
A;Title: Complementary DNA for the mouse homolog of the human amyloid beta protein precursor.
A;Reference number: A27485; MUID:88106489; PMID:3322280
A;Accession: A27485
A;Molecule type: mRNA
A;Residues: 1-695 <YAM>
A;Cross-references: GB:M18373; NID:g191568; PIDN:AAA37139.1; PID:g309085
A;Experimental source: brain
R;de Strooper, B.; van Leuven, F.; van den Berghe, H.
Biochim. Biophys. Acta 1129, 141-143, 1991
A;Title: The amyloid beta protein precursor or proteinase nexin II from mouse is closer related to its human homolog than previously reported.
A;Reference number: S19727; MUID:92096458; PMID:1756177
A;Accession: S19727
A;Molecule type: mRNA
A;Residues: 1-210,'G',212-220,'S',222-396,'A',398-402,'T',404-448,'A',450-695 <STR>
A;Cross-references: EMBL:X59379
R;Izumi, R.; Yamada, T.; Yoshikai, S.; Sasaki, H.; Hattori, M.; Sakaki, Y.
Gene 112, 189-195, 1992
A;Title: Positive and negative regulatory elements for the expression of the Alzheimer's disease amyloid precursor-encoding gene in mouse.
A;Reference number: I49485; MUID:92209998; PMID:1555768
A;Accession: I49485
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-19 <RES>
A;Cross-references: GB:D10603; NID:g220328; PIDN:BAA01456.1; PID:g220329
C;Genetics:
A;Map position: 16C3
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid; transmembrane protein

  Query Match 100.0%; Score 62; DB 2; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.0019;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKQYTSIHHG 11
      |||||||||||
Db 649 KKKQYTSIHHG 659

RESULT 4
S00550
Alzheimer's disease amyloid beta protein precursor - rat
N;Alternate names: beta-A4 amyloid protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C;Accession: S00550; A41245; A39820; S46251
R;Shivers, B.D.; Hilbich, C.; Multhaup, G.; Salbaum, M.; Beyreuther, K.; Seeburg, P.H.
EMBO J. 7, 1365-1370, 1988
A;Title: Alzheimer's disease amyloidogenic glycoprotein: expression pattern in rat brain suggests a role in cell contact.
A;Reference number: S00550; MUID:88312583; PMID:2900758
A;Accession: S00550

A;Molecule type: mRNA
A;Residues: 1-695 <SHI>
A;Cross-references: EMBL:X07648; NID:g55616; PIDN:CAA30488.1; PID:g55617
R;Schubert, D.; Schroeder, R.; LaCorbiere, M.; Saitoh, T.; Cole, G.
Science 241, 223-226, 1988
A;Title: Amyloid beta protein precursor is possibly a heparan sulfate proteoglycan core protein.
A;Reference number: A41245; MUID:88264430; PMID:2968652
A;Accession: A41245
A;Molecule type: protein
A;Residues: 18-37,'X',39-40,'X',42-44 <SCH>
A;Note: evidence for heparan sulfate attachment
R;Hesse, L.; Beher, D.; Masters, C.L.; Multhaup, G.
FEBS Lett. 349, 109-116, 1994
A;Title: The beta-A4 amyloid precursor protein binding to copper.
A;Reference number: S46251; MUID:94320627; PMID:7913895
A;Contents: annotation; copper binding sites
A;Note: rat peptides were isolated but not sequenced
R;Potempska, A.; Styles, J.; Mehta, P.; Kim, K.S.; Miller, D.L.
J. Biol. Chem. 266, 8464-8469, 1991
A;Title: Purification and tissue level of the beta-amyloid peptide precursor of rat brain.
A;Reference number: A39820; MUID:91217087; PMID:1673681
A;Accession: A39820
A;Status: preliminary
A;Molecule type: protein
A;Residues: 18-32 <POT>
A;Experimental source: brain
C;Comment: Deposition of amyloid protein as neurofibrillary tangles and/or plaques is characteristic of both Alzheimer's disease and Down's syndrome.
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid; glycoprotein; transmembrane protein
F;625-648/Domain: transmembrane #status predicted <TMM>

  Query Match 100.0%; Score 62; DB 2; Length 695;
  Best Local Similarity 100.0%; Pred. No. 0.0019;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        649 KKKQYTSIHHG 659
```

RESULT 5
QRHUA4
Alzheimer's disease amyloid beta protein precursor [validated] - human
N;Alternate names: Alzheimer's disease amyloid A4 protein; coagulation factor XIa inhibitor; proteinase nexin II (PN-II)
N;Contains: amyloid beta protein long, plaque form; amyloid beta protein short, vascular form; amyloid protein precursor splice form APP(695); amyloid protein precursor splice form APP(751); amyloid protein precursor splice form APP(770)
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1987 #sequence_revision 28-Jul-1995 #text_change 15-Sep-2000
C;Accession: S02260; S05194; A32277; A33260; A35486; I39452; I39451; I39453; I59562; A44017; B44017; A03134; A29030; A47584; A47585; S02638; S00707; S00925; A38949; A30320; B30320; C30320; A31087; A24668; A28583; A29302; A60805; JL0038; S06121; A60355; A59011; A38384; S29076; S38252; S32539; S48148; S48692; S51186; S51185; S51184; S51183; A54238; I58075; I52250; S09010; S10737; S24127; S43644
R;Lemaire, H.G.; Salbaum, J.M.; Multhaup, G.; Kang, J.; Bayney, R.M.; Unterbeck, A.; Beyreuther, K.; Mueller-Hill, B.
Nucleic Acids Res. 17, 517-522, 1989
A;Title: The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid is encoded by 16 exons.
A;Reference number: S02260; MUID:89128427; PMID:2783775
A;Accession: S02260
A;Molecule type: DNA
A;Residues: 1-288,'V',365-770 <LEM1>
A;Cross-references: EMBL:X13466
A;Note: alternative splice form APP(695)
R;Lemaire, H.G.
submitted to the EMBL Data Library, November 1988
A;Reference number: S05194
A;Accession: S05194
A;Molecule type: DNA
A;Residues: 1-14,'VW',17-288,'V',365-770 <LEM2>
A;Cross-references: EMBL:X13466; NID:g35598; PIDN:CAA31830.1; PID:g871360
A;Note: alternative splice form APP(695)
R;La Fauci, G.; Lahiri, D.K.; Salton, S.R.J.; Robakis, N.K.
Biochem. Biophys. Res. Commun. 159, 297-304, 1989
A;Title: Characterization of the 5'-end region and the first two exons of the beta-protein precursor gene.
A;Reference number: A32277; MUID:89165870; PMID:2538123
A;Accession: A32277
A;Molecule type: DNA
A;Residues: 1-75 <LAF>
A;Cross-references: GB:M24546; GB:M24547; NID:g341202; PIDN:AAC13654.1; PID:g516074
R;Johnstone, E.M.; Chaney, M.O.; Moore, R.E.; Ward, K.E.; Norris, F.H.; Little, S.P.
Biochem. Biophys. Res. Commun. 163, 1248-1255, 1989
A;Title: Alzheimer's disease amyloid peptide is encoded by two exons and shows similarity to soybean trypsin inhibitor.
A;Reference number: A33260; MUID:89392030; PMID:2675837
A;Accession: A33260
A;Molecule type: DNA
A;Residues: 656-737 <JOH>
A;Cross-references: GB:M29270; NID:g178863; PIDN:AAA51768.1; PID:g178865
R;Prelli, F.; Levy, E.; van Duinen, S.G.; Bots, G.T.A.M.; Luyendijk, W.; Frangione, B.
Biochem. Biophys. Res. Commun. 170, 301-307, 1990
A;Title: Expression of a normal and variant Alzheimer's beta-protein gene in amyloid of hereditary cerebral hemorrhage, Dutch type: DNA and protein diagnostic assays.
A;Reference number: A35486; MUID:90321244; PMID:2196878
A;Accession: A35486
A;Molecule type: DNA
A;Residues: 672-710 <PRE1>
A;Note: 693-Gln was found in DNA isolated from HCHWA-D patients
R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 87, 257-263, 1990
A;Title: Genomic organization of the human amyloid beta-protein precursor gene.
A;Reference number: I39451; MUID:90236318; PMID:2110105
A;Accession: I39452

A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-770 <YOS1>
A;Cross-references: GB:M33112; NID:g178613; PIDN:AAB59502.1; PID:g178616
A;Accession: I39451
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-530,'QWLMPVIPAFWEAKVGR' <YOS2>
A;Cross-references: GB:M34875; NID:g178608; PIDN:AAB59501.1; PID:g178615
R;Yoshikai, S.I.; Sasaki, H.; Doh-ura, K.; Furuya, H.; Sakaki, Y.
Gene 102, 291-292, 1991
A;Reference number: A59020; MUID:91340168; PMID:1908403
A;Contents: annotation; erratum
A;Note: revised physical map for reference I39451
R;Levy, E.; Carman, M.D.; Fernandez-Madrid, I.J.; Power, M.D.; Lieberburg, I.; van Duinen, S.G.; Bots, G.T.; Luyendijk, W.; Frangione, B.
Science 248, 1124-1126, 1990
A;Title: Mutation of the Alzheimer's disease amyloid gene in hereditary cerebral hemorrhage, Dutch type.
A;Reference number: I39453; MUID:90260663; PMID:2111584
A;Accession: I39453
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 656-737 <LEV>
A;Cross-references: GB:M37896; NID:g178618; PIDN:AAA51727.1; PID:g178620
A;Note: a mutation with 693-Gln is presented
R;Murrell, J.; Farlow, M.; Ghetti, B.; Benson, M.D.
Science 254, 97-99, 1991
A;Title: A mutation in the amyloid precursor protein associated with hereditary Alzheimer's disease.
A;Reference number: I59562; MUID:92022553; PMID:1925564
A;Accession: I59562
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 689-716,'F',718-737 <MUR>
A;Cross-references: GB:S57665; NID:g236720; PIDN:AAB19991.1; PID:g236721
R;Kamino, K.; Orr, H.T.; Payami, H.; Wijsman, E.M.; Alonso, M.E.; Pulst, S.M.; Anderson, L.; O'dahl, S.; Nemens, E.; White, J.A.; Sadovnick, A.D.; Ball, M.J.; Kaye, J.; Warren, A.; McInnis, M.; Antonarakis, S.E.; Korenberg, J.R.; Sharma, V.; Kukull, W.; Larson, E.; Heston, L.L.; Martin, G.M.; Bird, T.D.; Schellenberg, G.D.
Am. J. Hum. Genet. 51, 998-1014, 1992
A;Title: Linkage and mutational analysis of familial Alzheimer disease kindreds for the APP gene region.
A;Reference number: A44017; MUID:93035397; PMID:1415269
A;Accession: A44017
A;Molecule type: DNA
A;Residues: 687-692,'G',694-718 <KAM1>
A;Cross-references: GB:S45135; NID:g257377; PIDN:AAB23645.1; PID:g257378
A;Experimental source: familial Alzheimer disease family SB
A;Note: sequence extracted from NCBI backbone (NCBIP:115374)
A;Accession: B44017
A;Molecule type: DNA
A;Residues: 687-718 <KAM2>
A;Cross-references: GB:S45136; NID:g257379; PIDN:AAB23646.1; PID:g257380
A;Experimental source: familial Alzheimer disease family LIT
A;Note: sequence extracted from NCBI backbone (NCBIP:115376)
A;Note: this sequence has a silent mutation
R;Kang, J.; Lemaire, H.G.; Unterbeck, A.; Salbaum, J.M.; Masters, C.L.; Grzeschik, K.H.; Multhaup, G.; Beyreuther, K.; Muller-Hill, B.
Nature 325, 733-736, 1987
A;Title: The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor.
A;Reference number: A03134; MUID:87144572; PMID:2881207
A;Accession: A03134
A;Molecule type: mRNA
A;Residues: 1-288,'V',365-770 <KAN>
A;Cross-references: GB:Y00264; NID:g28525; PIDN:CAA68374.1; PID:g28526
A;Note: alternative splice form APP(695)
R;Robakis, N.K.; Ramakrishna, N.; Wolfe, G.; Wisniewski, H.M.
Proc. Natl. Acad. Sci. U.S.A. 84, 4190-4194, 1987
A;Title: Molecular cloning and characterization of a cDNA encoding the cerebrovascular and the neuritic plaque amyloid peptides.
A;Reference number: A29030; MUID:87231971; PMID:3035574
A;Accession: A29030
A;Molecule type: mRNA
A;Residues: 284-288,'V',365-646,'E',648-770 <ROB>
A;Cross-references: GB:M16765; NID:g178539; PIDN:AAA51722.1; PID:g178540
A;Note: the authors translated the codon GAG for residue 647 as Asp
R;Goldgaber, D.; Lerman, M.I.; McBride, O.W.; Saffiotti, U.; Gajdusek, D.C.
Science 235, 877-880, 1987
A;Title: Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease.
A;Reference number: A47584; MUID:87120328; PMID:3810169
A;Accession: A47584
A;Molecule type: mRNA
A;Residues: 674-756,'S',758-770 <GOL>
A;Cross-references: GB:M15533; NID:g178706; PIDN:AAA35540.1; PID:g178707
A;Experimental source: brain
R;Tanzi, R.E.; Gusella, J.F.; Watkins, P.C.; Bruns, G.A.P.; St George-Hyslop, P.; Van Keuren, M.L.; Patterson, D.; Pagan, S.; Kurnit, D.M.; Neve, R.L.
Science 235, 880-884, 1987
A;Title: Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus.
A;Reference number: A47585; MUID:87120329; PMID:2949367
A;Accession: A47585
A;Molecule type: mRNA
A;Residues: 674-703 <TAN1>
A;Cross-references: GB:M15532; NID:g177957; PIDN:AAA51564.1; PID:g177958
R;Dyrks, T.; Weidemann, A.; Multhaup, G.; Salbaum, J.M.; Lemaire, H.G.; Kang, J.; Mueller-Hill, B.; Masters, C.L.; Beyreuther, K.
EMBO J. 7, 949-957, 1988
A;Title: Identification, transmembrane orientation and biogenesis of the amyloid A4 precursor of Alzheimer's disease.
A;Reference number: S02638; MUID:88296437; PMID:2900137
A;Accession: S02638
A;Molecule type: mRNA
A;Residues: 672-678 <DYR>
R;Tanzi, R.E.; McClatchey, A.I.; Lamperti, E.D.; Villa-Komaroff, L.; Gusella, J.F.; Neve, R.L.
Nature 331, 528-530, 1988

A;Title: Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease.
A;Reference number: S00707; MUID:88122640; PMID:2893290
A;Accession: S00707
A;Molecule type: mRNA
A;Residues: 286-344,'I',365-366 <TAN2>
A;Cross-references: EMBL:X06982; NID:g28817; PIDN:CAA30042.1; PID:g929612
A;Experimental source: promyelocytic leukemia cell line HL60
A;Note: alternative splice form APP(751)
R;Ponte, P.; Gonzalez-DeWhitt, P.; Schilling, J.; Miller, J.; Hsu, D.; Greenberg, B.; Davis, K.; Wallace, W.; Lieberburg, I.; Fuller, F.; Cordell, B.
Nature 331, 525-527, 1988
A;Title: A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors.
A;Reference number: S00925; MUID:88122639; PMID:2893289
A;Accession: S00925
A;Molecule type: mRNA
A;Residues: 1-344,'I',365-770 <PO2>
A;Cross-references: GB:X06989; EMBL:Y00297; NID:g28720; PIDN:CAA30050.1; PID:g28721
A;Note: alternative splice form APP(751)
R;Kitaguchi, N.; Takahashi, Y.; Tokushima, Y.; Shiojiri, S.; Ito, H.
Nature 331, 530-532, 1988
A;Title: Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity.
A;Reference number: A38949; MUID:88122641; PMID:2893291
A;Accession: A38949
A;Molecule type: mRNA
A;Residues: 287-367 <KIT>
A;Cross-references: GB:X06981; NID:g28816; PIDN:CAA30041.1; PID:g929611
A;Experimental source: glioblastoma cell line
A;Note: alternative splice form APP(770)
R;Vitek, M.P.; Rasool, C.G.; de Sauvage, F.; Vitek, S.M.; Bartus, R.T.; Beer, B.; Ashton, R.A.; Macq, A.F.; Maloteaux, J.M.; Blume, A.J.; Octave, J.N.
Brain Res. Mol. Brain Res. 4, 121-131, 1988
A;Title: Absence of mutation in the beta-amyloid cDNAs cloned from the brains of three patients with sporadic Alzheimer's disease.
A;Reference number: A30320
A;Accession: A30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 284-288,'V',365-770 <VIT1>
A;Accession: B30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 122-288,'V',365-770 <VIT2>
A;Accession: C30320
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 606-770 <VIT3>
R;Zain, S.B.; Salim, M.; Chou, W.G.; Sajdel-Sulkowska, E.M.; Majocha, R.E.; Marotta, C.A.
Proc. Natl. Acad. Sci. U.S.A. 85, 929-933, 1988
A;Title: Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer disease brain: coding and noncoding regions of the fetal precursor mRNA are expressed in the cortex.
A;Reference number: A31087; MUID:88124954; PMID:2893379
A;Accession: A31087
A;Molecule type: mRNA
A;Residues: 507-770 <ZAI>
A;Cross-references: GB:M18734; NID:g178572; PIDN:AAA51726.1; PID:g178573
A;Note: the authors translated the codon GAA for residue 599 as Gly, ACC for residue 603 as Val, GTG for residue 604 as Glu, GAG for residue 605 as Leu, CTT for residue 607 as Pro, CCC for residue 608 as Val, GTG for residue 609 as Asn, AAT for residue 610 as Gly, and GGT for residue 655 as Ser
A;Note: the cited Genbank accession number, J03594, is not in release 101.0
R;Masters, C.L.; Multhaup, G.; Simms, G.; Pottgiesser, J.; Martins, R.N.; Beyreuther, K.

  Query Match 100.0%; Score 62; DB 1; Length 770;
  Best Local Similarity 100.0%; Pred. No. 0.0022;
  Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKQYTSIHHG 11
      |||||||||||
Db 724 KKKQYTSIHHG 734

RESULT 6
JH0773
Alzheimer's disease amyloid beta protein precursor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 13-Aug-1999
C;Accession: JH0773
R;Okado, H.; Okamoto, H.
Biochem. Biophys. Res. Commun. 189, 1561-1568, 1992
A;Title: A Xenopus homologue of the human beta-amyloid precursor protein: developmental regulation of its gene expression.
A;Reference number: JH0773; MUID:93129227; PMID:1282805
A;Accession: JH0773
A;Molecule type: mRNA
A;Residues: 1-747 <OKA>
A;Cross-references: GB:S52417; NID:g263150; PIDN:AAB24853.1; PID:g263151
A;Experimental source: larva
C;Superfamily: Alzheimer's disease amyloid beta protein; animal Kunitz-type proteinase inhibitor homology
C;Keywords: alternative splicing; amyloid
F;287-337/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>

  Query Match 95.2%; Score 59; DB 2; Length 747;
  Best Local Similarity 90.9%; Pred. No. 0.0075;
  Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKQYTSIHHG 11
      |||||:||||
Db 701 KKKQYTTIHHG 711

Search completed: October 4, 2004, 18:58:25
Job time : 18.2609 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:44:39 ; Search time 9.08696 Seconds
(without alignments)
63.032 Million cell updates/sec

Title: US-10-048-209-4
Perfect score: 62
Sequence: 1 KKKQYTSIHHG 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	751	1	A4_SAISC	Q95241 s amyloid b
2	62	100.0	770	1	A4_CAVPO	Q60495 c amyloid b
3	62	100.0	770	1	A4_HUMAN	P05067 h amyloid b
4	62	100.0	770	1	A4_MACFA	P53601 m amyloid b
5	62	100.0	770	1	A4_MOUSE	P12023 m amyloid b
6	62	100.0	770	1	A4_PIG	P79307 s amyloid b
7	62	100.0	770	1	A4_RAT	P08592 r amyloid b
8	59	95.2	737	1	A4_FUGRU	O93279 fugu rubrip
9	59	95.2	780	1	A4_TETFL	O73683 tetraodon f
10	43	69.4	379	1	ISPH_SYNY3	Q55643 synechocyst
11	39	62.9	452	1	NU4M_LUMTE	Q34949 lumbricus t
12	38	61.3	488	1	YGM9_YEAST	Q01163 saccharomyc
13	37	59.7	142	1	Y4KQ_RHISN	P55535 rhizobium s
14	37	59.7	653	1	APP1_MOUSE	Q03157 mus musculu
15	37	59.7	659	1	CYOB_BUCAP	Q8k994 buchnera ap
16	37	59.7	695	1	APP2_MOUSE	Q06335 mus musculu
17	37	59.7	763	1	APP2_HUMAN	Q06481 homo sapien
18	37	59.7	765	1	APP2_RAT	P15943 rattus norv
19	36	58.1	693	1	SC13_ARATH	Q9fq19 arabidopsis
20	36	58.1	749	1	YPRA_BACSU	P50830 bacillus su
21	36	58.1	1268	1	YRD6_CAEEL	Q09575 caenorhabdi
22	35	56.5	177	1	RELX_MESAU	Q64171 mesocricetu
23	35	56.5	197	1	PSD9_CAEEL	Q10920 caenorhabdi
24	35	56.5	211	1	14P_ARATH	O64628 arabidopsis
25	35	56.5	266	1	KLR8_MOUSE	Q60682 mus musculu
26	35	56.5	361	1	SERC_BACHD	Q9kdm4 bacillus ha
27	35	56.5	412	1	RBA1_CAEEL	P90917 caenorhabdi
28	35	56.5	443	1	GAT3_HUMAN	P23771 homo sapien
29	35	56.5	443	1	GAT3_MOUSE	P23772 mus musculu
30	35	56.5	497	1	DTPT_LACHE	O07380 lactobacill
31	35	56.5	516	1	Y067_MYCGE	P47313 mycoplasma
32	35	56.5	739	1	PURL_LISIN	Q92an9 listeria in
33	35	56.5	739	1	PURL_LISMO	Q8y6c1 listeria mo
34	35	56.5	753	1	CAT2_NEUCR	Q8x182 neurospora
35	35	56.5	782	1	MUS2_STAEP	Q8cp16 staphylococ
36	34	54.8	207	1	ATS5_BOVIN	Q9tt92 bos taurus
37	34	54.8	249	1	CREB_CHLVR	P51984 chlorohydra
38	34	54.8	348	1	NTRB_PROVU	P28788 proteus vul
39	34	54.8	366	1	TF3A_XENLA	P03001 xenopus lae
40	34	54.8	402	1	ISPH_ANASP	P58674 anabaena sp
41	34	54.8	580	1	SYD_MYCPE	Q8ewb7 mycoplasma
42	34	54.8	650	1	APP1_HUMAN	P51693 homo sapien
43	34	54.8	660	1	VNCS_PAVPN	P18547 porcine par
44	34	54.8	662	1	VNCS_PAVPK	P52502 porcine par
45	34	54.8	729	1	KEX1_YEAST	P09620 saccharomyc

ALIGNMENTS

RESULT 1
A4_SAISC
ID A4_SAISC STANDARD; PRT; 751 AA.
AC Q95241;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid
DE protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha); Soluble
DE APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-APP42);
DE Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40); Gamma-
DE CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE secretase C-terminal fragment 50); C31].
GN APP.
OS Saimiri sciureus (Common squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.
OX NCBI_TaxID=9521;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney, and Liver;
RX MEDLINE=96108492; PubMed=8532114;
RA Levy E., Amorim A., Frangione B., Walker L.C.;

```
RT   "Beta-amyloid precursor protein gene in squirrel monkeys with
RT   cerebral amyloid angiopathy.";
RL   Neurobiol. Aging 16:805-808(1995).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(0) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction. In vitro, copper-metallated APP induces neuronal
CC       death directly or is potentiated through Cu(II)-mediated low-
CC       density lipoprotein oxidation (By similarity). Can regulate
CC       neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=APP770;
CC         IsoId=Q95241-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=Q95241-2; Sequence=Not described;
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-720 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S81024; AAD14347.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
```

```
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Amyloid; Alternative splicing.
FT   SIGNAL        1     17       BY SIMILARITY.
FT   CHAIN        18    751       A4 PROTEIN.
FT   CHAIN        18    668       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN        18    652       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN       653    751       C99 (POTENTIAL).
FT   CHAIN       653    694       BETA-AMYLOID PROTEIN 42 (POTENTIAL).
FT   CHAIN       653    692       BETA-AMYLOID PROTEIN 40 (POTENTIAL).
FT   CHAIN       669    751       C83 (POTENTIAL).
FT   CHAIN       669    694       P3(42) (POTENTIAL).
FT   CHAIN       669    692       P3(40) (POTENTIAL).
FT   CHAIN       693    751       GAMMA-CTF(59) (POTENTIAL).
FT   CHAIN       695    751       GAMMA-CTF(57) (POTENTIAL).
FT   CHAIN       702    751       GAMMA-CTF(50) (POTENTIAL).
FT   CHAIN       721    751       C31 (POTENTIAL).
FT   DOMAIN       18    680       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    681    704       POTENTIAL.
FT   DOMAIN      705    751       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN       96    110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      181    188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN      291    341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN      316    344       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      363    428       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN      504    521       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN      713    732       INTERACTION WITH G(O)-ALPHA
FT                                (BY SIMILARITY).
FT   DOMAIN      230    260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      274    280       POLY-THR.
FT   SITE        144    144       REQUIRED FOR COPPER(II) REDUCTION
FT                                (BY SIMILARITY).
FT   ACT_SITE    301    302       REACTIVE BOND.
FT   SITE        652    653       CLEAVAGE (BY BETA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        653    654       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE        668    669       CLEAVAGE (BY ALPHA-SECRETASE)
FT                                (BY SIMILARITY).
FT   SITE        685    685       INVOLVED IN FREE RADICAL PROPAGATION
FT                                (BY SIMILARITY).
FT   SITE        687    687       INVOLVED IN OXIDATIVE REACTIONS
FT                                (BY SIMILARITY).
FT   SITE        692    693       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                                (BY SIMILARITY).
FT   SITE        694    695       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                                (BY SIMILARITY).
FT   SITE        701    702       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT                                (BY SIMILARITY).
FT   SITE        705    715       BASOLATERAL SORTING SIGNAL
FT                                (BY SIMILARITY).
FT   SITE        720    721       CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)
FT                                (BY SIMILARITY).
FT   SITE        738    741       ENDOCYTOSIS SIGNAL.
FT   SITE        740    743       NPXY MOTIF.

  Query Match             100.0%;  Score 62;  DB 1;  Length 751;
  Best Local Similarity   100.0%;  Pred. No. 0.0006;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        705 KKKQYTSIHHG 715


RESULT 2
A4_CAVPO
ID   A4_CAVPO       STANDARD;      PRT;   770 AA.
AC   Q60495; Q60496;
DT   10-OCT-2003 (Rel. 42, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); CTF-alpha; CTF-beta; Beta-amyloid
DE   protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); P3(42);
DE   P3(40); CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-
DE   CTF(57) (Gamma-secretase C-terminal fragment 57); C31].
GN   APP.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain, and Liver;
RX   MEDLINE=97236426; PubMed=9116031;
RA   Beck M., Mueller D., Bigl V.;
RT   "Amyloid precursor protein in Guinea pigs - complete cDNA sequence and
RT   alternative splicing.";
RL   Biochim. Biophys. Acta 1351:17-21(1997).
RN   [2]
RP   INTERACTION OF BETA-APP40 WITH APOE.
RX   MEDLINE=98007700; PubMed=9349544;
RA   Martel C.L., Mackic J.B., Matsubara E., Governale S., Miguel C.,
RA   Miao W., McComb J.G., Frangione B., Ghiso J., Zlokovic B.V.;
RT   "Isoform-specific effects of apolipoproteins E2, E3, and E4 on
RT   cerebral capillary sequestration and blood-brain barrier transport of
RT   circulating Alzheimer's amyloid beta.";
RL   J. Neurochem. 69:1995-2004(1997).
RN   [3]
RP   PROCESSING.
RX   MEDLINE=20084499; PubMed=10619481;
RA   Beck M., Brueckner M.K., Holzer M., Kaap S., Pannicke T., Arendt T.,
RA   Bigl V.;
```

```
RT   "Guinea-pig primary cell cultures provide a model to study expression
RT   and amyloidogenic processing of endogenous amyloid precursor
RT   protein.";
RL   Neuroscience 95:243-254(2000).
RN   [4]
RP   GAMMA-SECRETASE PROCESSING.
RX   MEDLINE=20576391; PubMed=11035007;
RA   Pinnix I., Musunuru U., Tun H., Sridharan A., Golde T., Eckman C.,
RA   Ziani-Cherif C., Onstead L., Sambamurti K.;
RT   "A novel gamma -secretase assay based on detection of the putative
RT   C-terminal fragment-gamma of amyloid beta protein precursor.";
RL   J. Biol. Chem. 276:481-487(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction (By similarity). In vitro, copper-metallated APP
CC       induces neuronal death directly or is potentiated through Cu(II)-
CC       mediated low-density lipoprotein oxidation (By similarity). Can
CC       regulate neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Beta-amyloid peptides bind to lipoproteins
CC       and apoliproteins E and J in the CSF and to HDL particles in
CC       plasma, inhibiting metal-catalyzed oxidation of lipoproteins.
CC   -!- FUNCTION: Appicans elicit adhesion of neural cells to the
CC       extracellular matrix and may regulate neurite outgrowth in the
CC       brain (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains), APPBP2 (via BaSS) and DDB1 (By similarity).
CC       Associates with microtubules in the presence of ATP and in a
CC       kinesin-dependent manner (By similarity). Soluble Abeta40 binds
CC       all three isoforms of APOE, in vitro and in vivo. When lipidated,
CC       ApoE3 appears to be the preferred amyloid binding isoform, while
CC       the apoE4 isoform-beta-APP40 complex is capable of being
CC       transported across the blood-brain barrier.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated pits
CC       (By similarity). During maturation, the immature APP (N-
CC       glycosylated in the endoplasmic reticulum) moves to the Golgi
CC       complex where complete maturation occurs (O-glycosylated and
CC       sulfated) (By similarity). After alpha-secretase cleavage, soluble
CC       APP is released into the extracellular space and the C-terminal is
CC       internalized to endosomes and lysosomes (By similarity). Some APP
CC       accumulates in secretory transport vesicles leaving the late Golgi
CC       compartment and returns to the cell surface (By similarity). APP
CC       sorts to the basolateral surface in epithelial cells (By
CC       similatity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms, missing exons 7,8 and 15, seem to
CC         exist. The L-isoforms, missing exon 15, are referred to as
CC         appicans;
CC       Name=APP770;
CC         IsoId=Q60495-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=Q60495-2; Sequence=VSP_007221, VSP_007222;
CC   -!- TISSUE SPECIFICITY: Isoform APP695 is the major isoform found in
CC       brain. The longer isoforms containing the BPTI domain are
CC       predominantly expressed in peripheral organs such as muscle and
CC       liver.
CC   -!- INDUCTION: Increased levels during neuronal differentiation.
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells.
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue (By similarity). The NPXY site is also involved in
CC       clathrin-mediated endocytosis.
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       CTF-alpha and CTF-beta. Subsequent processing of CTF-alpha by
CC       gamma-secretase yields P3 peptides. This is the major secretory
CC       pathway and is nonamyloidogenic. Alternatively,
CC       presenilin/nicastrin-mediated gamma-secretase processing of CTF-
CC       beta releases the amyloid beta proteins, amyloid-beta 40 (Abeta40)
CC       and amyloid-beta 42 (Abeta42), major components of amyloid
CC       plaques, and the corresponding cytotoxic C-terminal fragments
CC       (CTFs).
CC   -!- PTM: Proteolytically cleaved by caspase-3 during neuronal
CC       apoptosis (By similarity).
CC   -!- PTM: N- and O-glycosylated. O-linkage of chondroitin sulfate to
CC       the L-APP isoforms produces the APP proteoglycan core proteins,
CC       the appicans (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific (By similarity).
CC       Phosphorylation can affect APP processing, neuronal
CC       differentiation and interaction with other proteins.
CC   -!- PTM: Extracellular binding and reduction of copper, results in a
```

```
CC       corresponding oxidation of Cys-144 and Cys-158, and the formation
CC       of a disulfide bond (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates.
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X97631; CAA66230.1; -.
DR   EMBL; X99198; CAA67589.1; -.
DR   HSSP; P05067; 1BA4.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Alternative splicing; Amyloid.
FT   SIGNAL        1     17       BY SIMILARITY.
FT   CHAIN        18    770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN        18    687       SOLUBLE APP-ALPHA (BY SIMILARITY).
FT   CHAIN        18    671       SOLUBLE APP-BETA (BY SIMILARITY).
FT   CHAIN       672    770       CTF-ALPHA (BY SIMILARITY).
FT   CHAIN       672    713       BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
FT   CHAIN       672    711       BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
FT   CHAIN       688    770       CTF-BETA (BY SIMILARITY).
FT   CHAIN       688    713       P3(42) (BY SIMILARITY).
FT   CHAIN       688    711       P3(40) (BY SIMILARITY).
FT   CHAIN       712    770       GAMMA-CTF(59) (BY SIMILARITY).
FT   CHAIN       714    770       GAMMA-CTF(57) (BY SIMILARITY).

  Query Match           100.0%;  Score 62;  DB 1;  Length 770;
  Best Local Similarity  100.0%;  Pred. No. 0.00062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        724 KKKQYTSIHHG 734
```

```
RESULT 3
A4_HUMAN
ID   A4_HUMAN        STANDARD;      PRT;   770 AA.
AC   P05067; P09000; P78438; Q13764; Q13778; Q13793; Q16011; Q16014;
AC   Q16019; Q16020; Q9BT38; Q9UCA9; Q9UCB6; Q9UCC8; Q9UCD1; Q9UQ58;
DT   13-AUG-1987 (Rel. 05, Created)
DT   01-NOV-1991 (Rel. 20, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein) (Cerebral vascular amyloid peptide) (CVAP) (Protease
DE   nexin-II) (PN-II) (APPI) (PreA4) [Contains: Soluble APP-alpha (S-APP-
DE   alpha); Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42
DE   (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42);
DE   P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59)
DE   (Amyloid intracellular domain 59) (AID(59)); Gamma-CTF(57) (Gamma-
DE   secretase C-terminal fragment 57) (Amyloid intracellular domain 57)
DE   (AID(57)); Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50)
DE   (Amyloid intracellular domain 50) (AID(50)); C31].
GN   APP OR A4 OR AD1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=87144572; PubMed=2881207;
RA   Kang J., Lemaire H.-G., Unterbeck A., Salbaum J.M., Masters C.L.,
RA   Grzeschik K.-H., Multhaup G., Beyreuther K., Mueller-Hill B.;
RT   "The precursor of Alzheimer's disease amyloid A4 protein resembles a
RT   cell-surface receptor.";
RL   Nature 325:733-736(1987).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM APP751).
RC   TISSUE=Brain;
RX   MEDLINE=88122639; PubMed=2893289;
RA   Ponte P., Gonzalez-Dewhitt P., Schilling J., Miller J., Hsu D.,
RA   Greenberg B., Davis K., Wallace W., Lieberburg I., Fuller F.,
RA   Cordell B.;
RT   "A new A4 amyloid mRNA contains a domain homologous to serine
RT   proteinase inhibitors.";
RL   Nature 331:525-527(1988).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RX   MEDLINE=89128427; PubMed=2783775;
RA   Lemaire H.-G., Salbaum J.M., Multhaup G., Kang J., Bayney R.M.,
RA   Unterbeck A., Beyreuther K., Mueller-Hill B.;
RT   "The PreA4(695) precursor protein of Alzheimer's disease A4 amyloid
RT   is encoded by 16 exons.";
RL   Nucleic Acids Res. 17:517-522(1989).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM APP770).
RX   MEDLINE=90236318; PubMed=2110105;
RA   Yoshikai S.-I., Sasaki H., Doh-Ura K., Furuya H., Sakaki Y.;
RT   "Genomic organization of the human amyloid beta-protein precursor
RT   gene.";
```

RL Gene 87:257-263(1990).
RN [5]
RP ERRATUM, AND REVISIONS.
RA Yoshikai S.-I., Sasaki H., Doh-ura K., Furuya H., Sakaki Y.;
RL Gene 102:291-292(1991).
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM L-APP733).
RC TISSUE=Leukocyte;
RX MEDLINE=92268136; PubMed=1587857;
RA Koenig G., Moenning U., Czech C., Prior R., Banati R.,
RA Schreiter-Gasser U., Bauer J., Masters C.L., Beyreuther K.;
RT "Identification and differential expression of a novel alternative
RT splice isoform of the beta A4 amyloid precursor protein (APP) mRNA in
RT leukocytes and brain microglial cells.";
RL J. Biol. Chem. 267:10804-10809(1992).
RN [7]
RP SEQUENCE FROM N.A. (ISOFORM APP770).
RX MEDLINE=97263807; PubMed=9108164;
RA Hattori M., Tsukahara F., Furuhata Y., Tanahashi H., Hirose M.,
RA Saito M., Tsukuni S., Sakaki Y.;
RT "A novel method for making nested deletions and its application for
RT sequencing of a 300 kb region of human APP locus.";
RL Nucleic Acids Res. 25:1802-1808(1997).
RN [8]
RP SEQUENCE FROM N.A. (ISOFORM APP639).
RC TISSUE=Brain;
RX MEDLINE=22744650; PubMed=12859342;
RA Tang K., Wang C., Shen C., Sheng S., Ravid R., Jing N.;
RT "Identification of a novel alternative splicing isoform of human
RT amyloid precursor protein gene, APP639.";
RL Eur. J. Neurosci. 18:102-108(2003).
RN [9]
RP SEQUENCE FROM N.A. (ISOFORM APP305).
RC TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [10]
RP SEQUENCE OF 1-10 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89016647; PubMed=3140222;
RA Schon E.A., Mita S., Sadlock J., Herbert J.;
RT "A cDNA specifying the human amyloid beta precursor protein (ABPP)
RT encodes a 95-kDa polypeptide.";
RL Nucleic Acids Res. 16:9351-9351(1988).
RN [11]
RP ERRATUM, AND REVISIONS.
RA Mita S., Sadlock J., Herbert J., Schon E.A.;
RL Nucleic Acids Res. 16:11402-11402(1988).
RN [12]
RP SEQUENCE OF 1-75 FROM N.A.
RX MEDLINE=89165870; PubMed=2538123;
RA La Fauci G., Lahiri D.K., Salton S.R., Robakis N.K.;
RT "Characterization of the 5'-end region and the first two exons of the
RT beta-protein precursor gene.";
RL Biochem. Biophys. Res. Commun. 159:297-304(1989).
RN [13]
RP SEQUENCE OF 18-50.
RC TISSUE=Fibroblast;
RX MEDLINE=87250462; PubMed=3597385;
RA van Nostrand W.E., Cunningham D.D.;
RT "Purification of protease nexin II from human fibroblasts.";
RL J. Biol. Chem. 262:8508-8514(1987).
RN [14]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP751).
RC TISSUE=Brain;
RX MEDLINE=89346754; PubMed=2569763;
RA de Sauvage F., Octave J.N.;
RT "A novel mRNA of the A4 amyloid precursor gene coding for a possibly
RT secreted protein.";
RL Science 245:651-653(1989).
RN [15]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORM APP695).
RC TISSUE=Brain;
RX MEDLINE=87231971; PubMed=3035574;
RA Robakis N.K., Ramakrishna N., Wolfe G., Wisniewski H.M.;
RT "Molecular cloning and characterization of a cDNA encoding the
RT cerebrovascular and the neuritic plaque amyloid peptides.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:4190-4194(1987).
RN [16]
RP SEQUENCE OF 286-366 FROM N.A.
RX MEDLINE=88122640; PubMed=2893290;
RA Tanzi R.E., McClatchey A.I., Lamperti E.D., Villa-Komaroff L.,
RA Gusella J.F., Neve R.L.;
RT "Protease inhibitor domain encoded by an amyloid protein precursor
RT mRNA associated with Alzheimer's disease.";
RL Nature 331:528-530(1988).
RN [17]
RP SEQUENCE OF 287-367 FROM N.A.
RX MEDLINE=88122641; PubMed=2893291;
RA Kitaguchi N., Takahashi Y., Tokushima Y., Shiojiri S., Ito H.;
RT "Novel precursor of Alzheimer's disease amyloid protein shows
RT protease inhibitory activity.";
RL Nature 331:530-532(1988).
RN [18]
RP SEQUENCE OF 507-770 FROM N.A.

```
RC   TISSUE=Brain cortex;
RX   MEDLINE=88124954; PubMed=2893379;
RA   Zain S.B., Salim M., Chou W.G., Sajdel-Sulkowska E.M., Majocha R.E.,
RA   Marotta C.A.;
RT   "Molecular cloning of amyloid cDNA derived from mRNA of the Alzheimer
RT   disease brain: coding and noncoding regions of the fetal precursor
RT   mRNA are expressed in the cortex.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:929-933(1988).
RN   [19]
RP   SEQUENCE OF 523-555, AND COLLAGEN-BINDING DOMAIN.
RX   MEDLINE=96139497; PubMed=8576160;
RA   Beher D., Hesse L., Masters C.L., Multhaup G.;
RT   "Regulation of amyloid protein precursor (APP) binding to collagen and
RT   mapping of the binding sites on APP and collagen type I.";
RL   J. Biol. Chem. 271:1613-1620(1996).
RN   [20]
RP   SEQUENCE OF 655-737 FROM N.A., AND VARIANTS AD PHE-717; AD ILE-717
RP   AND AD GLY-717.
RX   MEDLINE=93236601; PubMed=8476439;
RA   Denman R.B., Rosenzcwaig R., Miller D.L.;
RT   "A system for studying the effect(s) of familial Alzheimer disease
RT   mutations on the processing of the beta-amyloid peptide precursor.";
RL   Biochem. Biophys. Res. Commun. 192:96-103(1993).
RN   [21]
RP   SEQUENCE OF 656-737 FROM N.A.
RX   MEDLINE=89392030; PubMed=2675837;
RA   Johnstone E.M., Chaney M.O., Moore R.E., Ward K.E., Norris F.H.,
RA   Little S.P.;
RT   "Alzheimer's disease amyloid peptide is encoded by two exons and shows
RT   similarity to soybean trypsin inhibitor.";
RL   Biochem. Biophys. Res. Commun. 163:1248-1255(1989).
RN   [22]

  Query Match             100.0%;  Score 62;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.00062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        724 KKKQYTSIHHG 734


RESULT 4
A4_MACFA
ID   A4_MACFA       STANDARD;      PRT;   770 AA.
AC   P53601; Q95KN7;
DT   01-OCT-1996 (Rel. 34, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-
DE   APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);
DE   Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
GN   APP.
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9541;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS APP695 AND APP770).
RC   TISSUE=Cerebellum;
RX   MEDLINE=91273117; PubMed=1905108;
RA   Podlisny M.B., Tolan D.R., Selkoe D.J.;
RT   "Homology of the amyloid beta protein precursor in monkey and human
RT   supports a primate model for beta amyloidosis in Alzheimer's
RT   disease.";
RL   Am. J. Pathol. 138:1423-1435(1991).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction. In vitro, copper-metallated APP induces neuronal
CC       death directly or is potentiated through Cu(II)-mediated low-
CC       density lipoprotein oxidation (By similarity). Can regulate
CC       neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity). The splice isoforms that contain the BPTI domain
CC       possess protease inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
```

```
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=Additional isoforms seem to exist;
CC       Name=APP770;
CC         IsoId=P53601-1; Sequence=Displayed;
CC       Name=APP695;
CC         IsoId=P53601-2; Sequence=VSP_000010, VSP_000011;
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M58727; AAA36829.1; -.
DR   EMBL; M58726; AAA36828.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Proteoglycan; Alternative splicing; Amyloid.
FT   SIGNAL       1    17       BY SIMILARITY.
FT   CHAIN       18   770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN       18   687       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN       18   671       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN      672   770       C99 (POTENTIAL).
FT   CHAIN      672   713       BETA-AMYLOID PROTEIN 42 (POTENTIAL).
FT   CHAIN      672   711       BETA-AMYLOID PROTEIN 40 (POTENTIAL).
FT   CHAIN      688   770       C83 (POTENTIAL).
FT   CHAIN      688   713       P3(42) (POTENTIAL).
FT   CHAIN      688   711       P3(40) (POTENTIAL).
FT   CHAIN      712   770       GAMMA-CTF(59) (POTENTIAL).
FT   CHAIN      714   770       GAMMA-CTF(57) (POTENTIAL).
FT   CHAIN      721   770       GAMMA-CTF(50) (POTENTIAL).
FT   CHAIN      740   770       C31 (POTENTIAL).
FT   DOMAIN      18   699       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   700   723       POTENTIAL.
FT   DOMAIN     724   770       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      96   110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     181   188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN     291   341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN     391   423       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     491   522       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     523   540       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN     732   751       INTERACTION WITH G(O)-ALPHA
FT                              (BY SIMILARITY).
FT   DOMAIN     230   260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN     274   280       POLY-THR.
FT   SITE       144   144       REQUIRED FOR COPPER(II) REDUCTION
FT                              (BY SIMILARITY).
FT   ACT_SITE   301   302       REACTIVE BOND (BY SIMILARITY).
FT   SITE       671   672       CLEAVAGE (BY BETA-SECRETASE)
FT                              (BY SIMILARITY).
```

```
FT   SITE        672    673       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE        687    688       CLEAVAGE (BY ALPHA-SECRETASE)
FT                                 (BY SIMILARITY).
FT   SITE        704    704       IMPLICATED IN FREE RADICAL PROPAGATION
FT                                 (BY SIMILARITY).
FT   SITE        706    706       INVOLVED IN OXIDATIVE REACTIONS
FT                                 (BY SIMILARITY).
FT   SITE        711    712       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                                 (BY SIMILARITY).
FT   SITE        713    714       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                                 (BY SIMILARITY).
FT   SITE        720    721       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)
FT                                 (BY SIMILARITY).
FT   SITE        724    734       BASOLATERAL SORTING SIGNAL
FT                                 (BY SIMILARITY).
FT   SITE        739    740       CLEAVAGE (BY CASPASES-3,-6,-8 OR -9)

  Query Match             100.0%;  Score 62;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.00062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        724 KKKQYTSIHHG 734


RESULT 5
A4_MOUSE
ID   A4_MOUSE       STANDARD;      PRT;   770 AA.
AC   P12023; P97487; P97942; Q99K32;
DT   01-OCT-1989 (Rel. 12, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) (Alzheimer's disease
DE   amyloid protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains:
DE   Soluble APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99
DE   (APP-C99); Beta-amyloid protein 42 (Beta-APP42); Beta-amyloid protein
DE   40 (Beta-APP40); C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase
DE   C-terminal fragment 59) (Amyloid intracellular domain 59) (AID(59))
DE   (APP-C59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57)
DE   (Amyloid intracellular domain 57) (AID(57)) (APP-C57); Gamma-CTF(50)
DE   (Gamma-secretase C-terminal fragment 50) (Amyloid intracellular domain
DE   50) (AID(50)); C31].
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=88106489; PubMed=3322280;
RA   Yamada T., Sasaki H., Furuya H., Miyata T., Goto I., Sakaki Y.;
RT   "Complementary DNA for the mouse homolog of the human amyloid beta
RT   protein precursor.";
RL   Biochem. Biophys. Res. Commun. 149:665-671(1987).
RN   [2]

RP   REVISIONS.
RA   Yamada T.;
RL   Submitted (MAR-1988) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=92096458; PubMed=1756177;
RA   de Strooper B., van Leuven F., van den Berghe H.;
RT   "The amyloid beta protein precursor or proteinase nexin II from mouse
RT   is closer related to its human homolog than previously reported.";
RL   Biochim. Biophys. Acta 1129:141-143(1991).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   STRAIN=SAMP8; TISSUE=Hippocampus;
RX   MEDLINE=21130647; PubMed=11235921;
RA   Kumar V.B., Vyas K., Franko M., Choudhary V., Buddhiraju C.,
RA   Alvarez J., Morley J.E.;
RT   "Molecular cloning, expression, and regulation of hippocampal amyloid
RT   precursor protein of senescence accelerated mouse (SAMP8).";
RL   Biochem. Cell Biol. 79:57-67(2001).
RN   [5]
RP   SEQUENCE OF 1-19 FROM N.A.
RX   MEDLINE=92209998; PubMed=1555768;
RA   Izumi R., Yamada T., Yoshikai S.I., Sasaki H., Hattori M.,
RA   Sakai Y.;
RT   "Positive and negative regulatory elements for the expression of the
RT   Alzheimer's disease amyloid precursor-encoding gene in mouse.";
RL   Gene 112:189-195(1992).
RN   [6]
RP   PARTIAL SEQUENCE FROM N.A. (ISOFORM APP770).
RC   TISSUE=Breast tumor;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   SEQUENCE OF 281-380 FROM N.A., AND ALTERNATIVE SPLICING.
RC   TISSUE=Brain, and Kidney;
```

RX MEDLINE=89149813; PubMed=2493250;
RA Yamada T., Sasaki H., Dohura K., Goto I., Sakaki Y.;
RT "Structure and expression of the alternatively-spliced forms of mRNA
RT for the mouse homolog of Alzheimer's disease amyloid beta protein
RT precursor.";
RL Biochem. Biophys. Res. Commun. 158:906-912(1989).
RN [8]
RP SEQUENCE OF 289-364 FROM N.A.
RC STRAIN=CD-1; TISSUE=Placenta;
RX MEDLINE=89345111; PubMed=2569710;
RA Fukuchi K., Martin G.M., Deeb S.S.;
RT "Sequence of the protease inhibitor domain of the A4 amyloid protein
RT precursor of Mus domesticus.";
RL Nucleic Acids Res. 17:5396-5396(1989).
RN [9]
RP SEQUENCE OF 656-737 FROM N.A.
RC STRAIN=129/Sv;
RA Wragg M.A., Busfield F., Duff K., Korenblat K., Capecchi M.,
RA Loring J.F., Goate A.M.;
RT "Introduction of six mutations into the mouse genome using 'Hit and
RT Run' gene-targeting: introduction of familial Alzheimer's disease
RT mutations into the mouse amyloid precursor protein gene and
RT humanization of the A-beta fragment.";
RL Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
RN [10]
RP TISSUE SPECIFICITY OF ALTERNATIVE SPLICED FORMS.
RX MEDLINE=93287808; PubMed=8510506;
RA Sola C., Mengod G., Ghetti B., Palacios J.M., Triarhou L.C.;
RT "Regional distribution of the alternatively spliced isoforms of beta
RT APP RNA transcript in the brain of normal, heterozygous and
RT homozygous weaver mutant mice as revealed by in situ hybridization
RT histochemistry.";
RL Brain Res. Mol. Brain Res. 17:340-346(1993).
RN [11]
RP INTERACTION WITH KNS2.
RX MEDLINE=21010507; PubMed=11144355;
RA Kamal A., Stokin G.B., Yang Z., Xia C.-H., Goldstein L.S.;
RT "Axonal transport of amyloid precursor protein is mediated by direct
RT binding to the kinesin light chain subunit of kinesin-I.";
RL Neuron 28:449-459(2000).
RN [12]
RP C-TERMINAL PROTEIN-PROTEIN INTERACTIONS, AND MUTAGENESIS OF TYR-728;
RP THR-743; TYR-757; ASN-759 AND TYR-762.
RX MEDLINE=21408156; PubMed=11517249;
RA Matsuda S., Yasukawa T., Homma Y., Ito Y., Niikura T., Hiraki T.,
RA Hirai S., Ohno S., Kita Y., Kawasumi M., Kouyama K., Yamamoto T.,
RA Kyriakis J.M., Nishimoto I.;
RT "C-jun N-terminal kinase (JNK)-interacting protein-1b/islet-brain-1
RT scaffolds Alzheimer's amyloid precursor protein with JNK.";
RL J. Neurosci. 21:6597-6607(2001).
RN [13]
RP INTERACTION WITH MAPK8IP1, AND PHOSPHORYLATION.
RX MEDLINE=22028091; PubMed=11912189;
RA Taru H., Iijima K.-I., Hase M., Kirino Y., Yagi Y., Suzuki T.;
RT "Interaction of Alzheimer's beta-amyloid precursor family proteins
RT with scaffold proteins of the JNK signaling cascade.";
RL J. Biol. Chem. 277:20070-20078(2002).
RN [14]
RP INTERACTION OF CTF PEPTIDES WITH NUMB.
RX MEDLINE=22008109; PubMed=12011466;
RA Roncarati R., Sestan N., Scheinfeld M.H., Berechid B.E., Lopez P.A.,
RA Meucci O., McGlade J.C., Rakic P., D'Adamio L.;
RT "The gamma-secretase-generated intracellular domain of beta-amyloid
RT precursor protein binds Numb and inhibits Notch signaling.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:7102-7107(2002).
RN [15]
RP GAMMA-SECRETASE PROCESSING, AND INTERACTION WITH APBB1.
RX MEDLINE=21437805; PubMed=11553691;
RA Cupers P., Orlans I., Craessaerts K., Annaert W., De Strooper B.;
RT "The amyloid precursor protein (APP)-cytoplasmic fragment generated by
RT gamma-secretase is rapidly degraded but distributes partially in a
RT nuclear fraction of neurones in culture.";
RL J. Neurochem. 78:1168-1178(2001).
CC -!- FUNCTION: Functions as a cell surface receptor and performs
CC physiological functions on the surface of neurons relevant to
CC neurite growth, neuronal adhesion and axonogenesis. Involved in
CC cell mobility and transcription regulation through protein-protein
CC interactions. Can promote transcription activation through binding
CC to APBB1/Tip60 and inhibit Notch signaling through interaction
CC with Numb. Couples to apoptosis-inducing pathways such as those
CC mediated by G(O) and JIP. Inhibits G(O) alpha ATPase activity (By
CC similarity). Acts as a kinesin I membrane receptor, mediating the
CC axonal transport of beta-secretase and presenilin 1. May be
CC involved in copper homeostasis/oxidative stress through copper ion
CC reduction. Can regulate neurite outgrowth through binding to
CC components of the extracellular matrix such as heparin and
CC collagen I and IV (By similarity). The splice isoforms that
CC contain the BPTI domain possess protease inhibitor activity (By
CC similarity).
CC -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC with metal-reducing activity. Bind transient metals such as
CC copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
CC only weakly transient metals and have little reducing activity due
CC to substitutions of transient metal chelating residues. Beta-APP42
CC may activate mononuclear phagocytes in the brain and elicit
CC inflammatory responses. Promotes both tau aggregation and TPK II-
CC mediated phosphorylation (By similarity).
CC -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC peptides, including C31, are potent enhancers of neuronal
CC apoptosis.
CC -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC cytoplasmic proteins, including APBB family members, the APBA
CC family, MAPK8IP1, SHC1, Numb and Dab1. Binding to Dab1 inhibits
CC its serine phosphorylation. Also interacts with GPCR-like protein
CC BPP, FPRL1, APPBP1, IB1, KNS2 (via its TPR domains), APPBP2 (via
CC BaSS) and DDB1 (By similarity). In vitro, it binds MAPT via the
CC MT-binding domains (By similarity). Associates with microtubules
CC in the presence of ATP and in a kinesin-dependent manner (By
CC similarity). Interacts, through a C-terminal domain, with GNAO1
CC (By similarity). Amyloid beta-42 binds CHRNA7 in hippocampal
CC neurons (By similarity). Beta-amyloid associates with HADH2 (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC protein that rapidly becomes internalized via clathrin-coated

```
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete

  Query Match             100.0%;  Score 62;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.00062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        724 KKKQYTSIHHG 734


RESULT 6
A4_PIG
ID   A4_PIG         STANDARD;      PRT;   770 AA.
AC   P79307; Q29023; Q9TUI0;
DT   01-NOV-1997 (Rel. 35, Created)
DT   10-OCT-2003 (Rel. 42, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (ABPP) [Alzheimer's disease
DE   amyloid protein homolog) [Contains: Soluble APP-alpha (S-APP-alpha);
DE   Soluble APP-beta (S-APP-beta); C99; Beta-amyloid protein 42 (Beta-
DE   APP42); Beta-amyloid protein 40 (Beta-APP40); C83; P3(42); P3(40);
DE   Gamma-CTF(59) (Gamma-secretase C-terminal fragment 59); Gamma-CTF(57)
DE   (Gamma-secretase C-terminal fragment 57); Gamma-CTF(50) (Gamma-
DE   secretase C-terminal fragment 50); C31].
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kimura A., Takahashi T.;
RT   "Amyloid precursor protein 770.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 1-136 FROM N.A.
RC   TISSUE=Small intestine;
RA   Winteroe A.K., Fredholm M.;
RT   "Evaluation and characterization of a porcine small intestine cDNA
RT   library.";
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 667-723 FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=92017079; PubMed=1656157;
RA   Johnstone E.M., Chaney M.O., Norris F.H., Pascual R., Little S.P.;
RT   "Conservation of the sequence of the Alzheimer's disease amyloid
RT   peptide in dog, polar bear and five other mammals by cross-species
RT   polymerase chain reaction analysis.";
RL   Brain Res. Mol. Brain Res. 10:299-305(1991).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP (By
CC       similarity). Inhibits G(O) alpha ATPase activity (By similarity).
CC       Acts as a kinesin I membrane receptor, mediating the axonal
CC       transport of beta-secretase and presenilin 1 (By similarity). May
CC       be involved in copper homeostasis/oxidative stress through copper
CC       ion reduction (By similarity). In vitro, copper-metallated APP
CC       induces neuronal death directly or is potentiated through Cu(II)-
CC       mediated low-density lipoprotein oxidation (By similarity). Can
CC       regulate neurite outgrowth through binding to components of the
CC       extracellular matrix such as heparin and collagen I and IV (By
CC       similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron (By similarity).
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, and SHC1, Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains) (By similarity), APPBP2 (via BaSS) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the
CC       endoplasmic reticulum) moves to the Golgi complex where complete
CC       maturation occurs (O-glycosylated and sulfated). After alpha-
CC       secretase cleavage, soluble APP is released into the extracellular
CC       space and the C-terminal is internalized to endosomes and
CC       lysosomes. Some APP accumulates in secretory transport vesicles
CC       leaving the late Golgi compartment and returns to the cell
CC       surface. Gamma-CTF(59) peptide is located to both the cytoplasm
CC       and nuclei of neurons (By similarity).
CC   -!- DOMAIN: The basolateral sorting signal (BaSS) is required for
CC       sorting of membrane proteins to the basolateral surface of
CC       epithelial cells (By similarity).
CC   -!- DOMAIN: The NPXY sequence motif found in many tyrosine-
CC       phosphorylated proteins is required for the specific binding of
CC       the PID domain. However additional amino acids either N- or C-
CC       terminal to the NPXY motif are often required for complete
CC       interaction. The PID domain-containing proteins which bind APP
CC       require the YENPTY motif for full interaction. These interactions
CC       are independent of phosphorylation on the terminal tyrosine
CC       residue. The NPXY site is also involved in clathrin-mediated
CC       endocytosis (By similarity).
CC   -!- PTM: Proteolytically processed under normal cellular conditions.
CC       Cleavage by alpha-secretase or alternatively by beta-secretase
CC       leads to generation and extracellular release of soluble APP
CC       peptides, S-APP-alpha and S-APP-beta, respectively, and the
CC       retention of corresponding membrane-anchored C-terminal fragments,
CC       C83 and C99. Subsequent processing of C83 by gamma-secretase
CC       yields P3 peptides. This is the major secretory pathway and is
```

```
CC       nonamyloidogenic. Alternatively, presenilin/nicastrin-mediated
CC       gamma-secretase processing of C99 releases the amyloid beta
CC       proteins, amyloid-beta 40 (Abeta40) and amyloid-beta 42 (Abeta42),
CC       major components of amyloid plaques, and the cytotoxic C-terminal
CC       fragments, gamma-CTF(50), gamma-CTF(57) and gamma-CTF(59) (By
CC       similarity).
CC   -!- PTM: Proteolytically cleaved by caspases during neuronal apoptosis
CC       (By similarity). Cleavage at Asp-739 by either caspase-3, -8 or -9
CC       results in the production of the neurotoxic C31 peptide and the
CC       increased production of beta-amyloid peptides (By similarity).
CC   -!- PTM: N- and O-glycosylated (By similarity).
CC   -!- PTM: Phosphorylation in the C-terminal on tyrosine, threonine and
CC       serine residues is neuron-specific. Phosphorylation can affect APP
CC       processing, neuronal differentiation and interaction with other
CC       proteins (By similarity).
CC   -!- PTM: Extracellular binding and reduction of copper, results in a
CC       corresponding oxidation of Cys-144 and Cys-158, and the formation
CC       of a disulfide bond (By similarity).
CC   -!- MISCELLANEOUS: Chelation of metal ions, notably copper, iron and
CC       zinc, can induce histidine-bridging between beta-amyloid molecules
CC       resulting in beta-amyloid-metal aggregates (By similarity).
CC       Extracellular zinc-binding increases binding of heparin to APP and
CC       inhibits collagen-binding (By similarity).
CC   -!- SIMILARITY: Belongs to the APP family.
CC   -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB032550; BAA84580.1; -.
DR   EMBL; Z84022; CAB06313.1; -.
DR   EMBL; X56127; CAA39592.1; -.
DR   HSSP; P05067; 1AAP.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Apoptosis; Endocytosis; Cell adhesion; Serine protease inhibitor;
KW   Coated pits; Neurone; Heparin-binding; Metal-binding; Copper; Iron;
KW   Zinc; Signal; Transmembrane; Glycoprotein; Phosphorylation;
KW   Amyloid.
FT   SIGNAL       1     17       BY SIMILARITY.
FT   CHAIN       18    770       AMYLOID BETA A4 PROTEIN.
FT   CHAIN       18    687       SOLUBLE APP-ALPHA (POTENTIAL).
FT   CHAIN       18    671       SOLUBLE APP-BETA (POTENTIAL).
FT   CHAIN      672    770       C99 (BY SIMILARITY).
FT   CHAIN      672    713       BETA-AMYLOID PROTEIN 42 (BY SIMILARITY).
FT   CHAIN      672    711       BETA-AMYLOID PROTEIN 40 (BY SIMILARITY).
FT   CHAIN      688    770       C83 (BY SIMILARITY).
FT   CHAIN      688    713       P3(42) (BY SIMILARITY).
FT   CHAIN      688    711       P3(40) (BY SIMILARITY).
FT   CHAIN      712    770       GAMMA-CTF(59).
FT   CHAIN      714    770       GAMMA-CTF(57).
FT   CHAIN      721    770       GAMMA-CTF(50) (BY SIMILARITY).
FT   CHAIN      740    770       C31 (DURING APOPTOSIS) (BY SIMILARITY).
FT   DOMAIN      18    699       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   700    723       POTENTIAL.
FT   DOMAIN     724    770       CYTOPLASMIC (POTENTIAL).
FT   DOMAIN      96    110       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     135    155       COPPER-BINDING (BY SIMILARITY).
FT   DOMAIN     181    188       ZINC-BINDING (BY SIMILARITY).
FT   DOMAIN     291    341       BPTI/KUNITZ INHIBITOR.
FT   DOMAIN     391    423       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     491    522       HEPARIN-BINDING (BY SIMILARITY).
FT   DOMAIN     523    540       COLLAGEN-BINDING (BY SIMILARITY).
FT   DOMAIN     732    751       INTERACTION WITH G(O)-ALPHA (BY
FT                               SIMILARITY).
FT   DOMAIN     230    260       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN     274    280       POLY-THR.
FT   SITE       144    144       REQUIRED FOR COPPER(II) REDUCTION
FT                               (BY SIMILARITY).
FT   ACT_SITE   301    302       REACTIVE BOND (BY SIMILARITY).
FT   SITE       671    672       CLEAVAGE (BY BETA-SECRETASE)
FT                               (BY SIMILARITY).
FT   SITE       672    673       CLEAVAGE (BY CASPASE-6) (BY SIMILARITY).
FT   SITE       687    688       CLEAVAGE (BY ALPHA-SECRETASE)
FT                               (BY SIMILARITY).
FT   SITE       704    704       IMPLICATED IN FREE RADICAL PROPAGATION
FT                               (BY SIMILARITY).
FT   SITE       706    706       INVOLVED IN OXIDATIVE REACTIONS
FT                               (BY SIMILARITY).
FT   SITE       711    712       CLEAVAGE (BY GAMMA-SECRETASE; SITE 1)
FT                               (BY SIMILARITY).
FT   SITE       713    714       CLEAVAGE (BY GAMMA-SECRETASE; SITE 2)
FT                               (BY SIMILARITY).
FT   SITE       720    721       CLEAVAGE (BY GAMMA-SECRETASE; SITE 3)

  Query Match             100.0%;  Score 62;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.00062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        724 KKKQYTSIHHG 734


RESULT 7
A4_RAT
ID   A4_RAT          STANDARD;      PRT;   770 AA.
AC   P08592;
```

```
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Amyloid beta A4 protein precursor (APP) (Alzheimer's disease amyloid
DE   protein homolog) (Amyloidogenic glycoprotein) (AG) [Contains: Soluble
DE   APP-alpha (S-APP-alpha); Soluble APP-beta (S-APP-beta); C99; Beta-
DE   amyloid protein 42 (Beta-APP42); Beta-amyloid protein 40 (Beta-APP40);
DE   C83; P3(42); P3(40); Gamma-CTF(59) (Gamma-secretase C-terminal
DE   fragment 59); Gamma-CTF(57) (Gamma-secretase C-terminal fragment 57);
DE   Gamma-CTF(50) (Gamma-secretase C-terminal fragment 50); C31].
GN   APP.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM APP695).
RC   TISSUE=Brain;
RX   MEDLINE=88312583; PubMed=2900758;
RA   Shivers B.D., Hilbich C., Multhaup G., Salbaum J.M., Beyreuther K.,
RA   Seeburg P.H.;
RT   "Alzheimer's disease amyloidogenic glycoprotein: expression pattern
RT   in rat brain suggests a role in cell contact.";
RL   EMBO J. 7:1365-1370(1988).
RN   [2]
RP   SEQUENCE OF 289-364 FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=89183625; PubMed=2648331;
RA   Kang J., Mueller-Hill B.;
RT   "The sequence of the two extra exons in rat preA4.";
RL   Nucleic Acids Res. 17:2130-2130(1989).
RN   [3]
RP   SEQUENCE OF 720-730, AND MASS SPECTROMETRY.
RX   MEDLINE=21443797; PubMed=11483588;
RA   Gu Y., Misonou H., Sato T., Dohmae N., Takio K., Ihara Y.;
RT   "Distinct intramembrane cleavage of the beta-amyloid precursor protein
RT   family resembling gamma-secretase-like cleavage of Notch.";
RL   J. Biol. Chem. 276:35235-35238(2001).
RN   [4]
RP   ALTERNATIVE SPLICING.
RX   MEDLINE=96187032; PubMed=8624099;
RA   Sandbrink R., Masters C.L., Beyreuther K.;
RT   "APP gene family. Alternative splicing generates functionally related
RT   isoforms.";
RL   Ann. N.Y. Acad. Sci. 777:281-287(1996).
RN   [5]
RP   TISSUE SPECIFICITY OF APPICAN.
RX   MEDLINE=95263526; PubMed=7744833;
RA   Shioi J., Pangalos M.N., Ripellino J.A., Vassilacopoulou D.,
RA   Mytilineou C., Margolis R.U., Robakis N.K.;
RT   "The Alzheimer amyloid precursor proteoglycan (appican) is present in
RT   brain and is produced by astrocytes but not by neurons in primary
RT   neural cultures.";
RL   J. Biol. Chem. 270:11839-11844(1995).
RN   [6]
RP   TISSUE SPECIFICITY OF ISOFORMS.
RX   MEDLINE=97150061; PubMed=8996834;
RA   Sandbrink R., Monning U., Masters C.L., Beyreuther K.;
RT   "Expression of the APP gene family in brain cells, brain development
RT   and aging.";
RL   Gerontology 43:119-131(1997).
RN   [7]
RP   INTERACTION WITH DDB1, AND MUTAGENESIS OF TYR-757; ASN-759 AND
RP   TYR-762.
RX   MEDLINE=99127916; PubMed=9930726;
RA   Watanabe T., Sukegawa J., Tomita S., Iijima K.-I., Oguchi S.,
RA   Suzuki T., Nairn A.C., Greengard P.;
RT   "A 127-kDa protein (UV-DDB) binds to the cytoplasmic domain of the
RT   Alzheimer's amyloid precursor protein.";
RL   J. Neurochem. 72:549-556(1999).
RN   [8]
RP   INTERACTION WITH GNAO1, AND MUTAGENESIS OF 732-HIS-HIS-733.
RX   MEDLINE=99162676; PubMed=10024358;
RA   Brouillet E., Trembleau A., Galanaud D., Volovitch M., Bouillot C.,
RA   Valenza C., Prochiantz A., Allinquant B.;
RT   "The amyloid precursor protein interacts with Go heterotrimeric
RT   protein within a cell compartment specialized in signal
RT   transduction.";
RL   J. Neurosci. 19:1717-1727(1999).
RN   [9]
RP   CHARACTERISTICS OF APPICAN, AND MUTAGENESIS OF SER-656.
RX   MEDLINE=95256193; PubMed=7737970;
RA   Pangalos M.N., Efthimiopoulos S., Shioi J., Robakis N.K.;
RT   "The chondroitin sulfate attachment site of appican is formed by
RT   splicing out exon 15 of the amyloid precursor gene.";
RL   J. Biol. Chem. 270:10388-10391(1995).
RN   [10]
RP   BETA-AMYLOID METAL-BINDING.
RX   MEDLINE=99316162; PubMed=10386999;
RA   Huang X., Atwood C.S., Hartshorn M.A., Multhaup G., Goldstein L.E.,
RA   Scarpa R.C., Cuajungco M.P., Gray D.N., Lim J., Moir R.D., Tanzi R.E.,
RA   Bush A.I.;
RT   "The A beta peptide of Alzheimer's disease directly produces hydrogen
RT   peroxide through metal ion reduction.";
RL   Biochemistry 38:7609-7616(1999).
RN   [11]
RP   BETA-AMYLOID ZINC BINDING.
RX   MEDLINE=99343552; PubMed=10413512;
RA   Liu S.T., Howlett G., Barrow C.J.;
RT   "Histidine-13 is a crucial residue in the zinc ion-induced aggregation
RT   of the A beta peptide of Alzheimer's disease.";
RL   Biochemistry 38:9373-9378(1999).
RN   [12]
RP   IMPORTANCE OF GLY-704 IN FREE RADICAL PROPAGATION, AND MUTAGENESIS OF
RP   GLY-704.
RX   MEDLINE=21956095; PubMed=11959460;
RA   Kanski J., Varadarajan S., Aksenova M., Butterfield D.A.;
RT   "Role of glycine-33 and methionine-35 in Alzheimer's amyloid beta-
RT   peptide 1-42-associated oxidative stress and neurotoxicity.";
RL   Biochim. Biophys. Acta 1586:190-198(2001).
RN   [13]
RP   PHOSPHORYLATION.
RX   MEDLINE=97239592; PubMed=9085254;
RA   Oishi M., Nairn A.C., Czernik A.J., Lim G.S., Isohara T., Gandy S.E.,
```

```
RA   Greengard P., Suzuki T.;
RT   "The cytoplasmic domain of Alzheimer's amyloid precursor protein is
RT   phosphorylated at Thr654, Ser655, and Thr668 in adult rat brain and
RT   cultured cells.";
RL   Mol. Med. 3:111-123(1997).
RN   [14]
RP   PHOSPHORYLATION ON SER-730.
RX   MEDLINE=99262094; PubMed=10329382;
RA   Isohara T., Horiuchi A., Watanabe T., Ando K., Czernik A.J., Uno I.,
RA   Greengard P., Nairn A.C., Suzuki T.;
RT   "Phosphorylation of the cytoplasmic domain of Alzheimer's beta-amyloid
RT   precursor protein at Ser655 by a novel protein kinase.";
RL   Biochem. Biophys. Res. Commun. 258:300-305(1999).
RN   [15]
RP   PHOSPHORYLATION, INDUCTION, SUBCELLULAR LOCATION, AND MUTAGENESIS OF
RP   THR-743.
RX   MEDLINE=99274744; PubMed=10341243;
RA   Ando K., Oishi M., Takeda S., Iijima K.-I., Isohara T., Nairn A.C.,
RA   Kirino Y., Greengard P., Suzuki T.;
RT   "Role of phosphorylation of Alzheimer's amyloid precursor protein
RT   during neuronal differentiation.";
RL   J. Neurosci. 19:4421-4427(1999).
RN   [16]
RP   PHOSPHORYLATION ON THR-743.
RX   MEDLINE=20396183; PubMed=10936190;
RA   Iijima K.-I., Ando K., Takeda S., Satoh Y., Seki T., Itohara S.,
RA   Greengard P., Kirino Y., Nairn A.C., Suzuki T.;
RT   "Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor
RT   protein by cyclin-dependent kinase 5.";
RL   J. Neurochem. 75:1085-1091(2000).
RN   [17]
RP   CARBOHYDRATE STRUCTURE OF APPICAN.
RX   MEDLINE=21463085; PubMed=11479316;
RA   Tsuchida K., Shioi J., Yamada S., Boghosian G., Wu A., Cai H.,
RA   Sugahara K., Robakis N.K.;
RT   "Appican, the proteoglycan form of the amyloid precursor protein,
RT   contains chondroitin sulfate E in the repeating disaccharide region
RT   and 4-O-sulfated galactose in the linkage region.";
RL   J. Biol. Chem. 276:37155-37160(2001).
CC   -!- FUNCTION: Functions as a cell surface receptor and performs
CC       physiological functions on the surface of neurons relevant to
CC       neurite growth, neuronal adhesion and axonogenesis. Involved in
CC       cell mobility and transcription regulation through protein-protein
CC       interactions (By similarity). Can promote transcription activation
CC       through binding to APBB1/Tip60 and inhibit Notch signaling through
CC       interaction with Numb (By similarity). Couples to apoptosis-
CC       inducing pathways such as those mediated by G(O) and JIP. Inhibits
CC       G(O) alpha ATPase activity. Acts as a kinesin I membrane receptor,
CC       mediating the axonal transport of beta-secretase and presenilin 1
CC       (By similarity). May be involved in copper homeostasis/oxidative
CC       stress through copper ion reduction. Can regulate neurite
CC       outgrowth through binding to components of the extracellular
CC       matrix such as heparin and collagen I and IV (By similarity). The
CC       splice isoforms that contain the BPTI domain possess protease
CC       inhibitor activity (By similarity).
CC   -!- FUNCTION: Beta-amyloid peptides are lipophilic metal chelators
CC       with metal-reducing activity. Bind transient metals such as
CC       copper, zinc and iron. Rat and mouse beta-amyloid peptides bind
CC       only weakly transient metals and have little reducing activity due
CC       to substitutions of transient metal chelating residues. Beta-APP42
CC       may activate mononuclear phagocytes in the brain and elicit
CC       inflammatory responses. Promotes both tau aggregation and TPK II-
CC       mediated phosphorylation (By similarity).
CC   -!- FUNCTION: Appicans elicit adhesion of neural cells to the
CC       extracellular matrix and may regulate neurite outgrowth in the
CC       brain.
CC   -!- FUNCTION: The gamma-CTF peptides as well as the caspase-cleaved
CC       peptides, including C31, are potent enhancers of neuronal
CC       apoptosis (By similarity).
CC   -!- SUBUNIT: Binds, via its C-terminal, to the PID domain of several
CC       cytoplasmic proteins, including APBB family members, the APBA
CC       family, MAPK8IP1, SHC1 and Numb and Dab1 (By similarity). Binding
CC       to Dab1 inhibits its serine phosphorylation (By similarity). Also
CC       interacts with GPCR-like protein BPP, FPRL1, APPBP1, IB1, KNS2
CC       (via its TPR domains), APPBP2 (via BaSS) (By similarity) and DDB1.
CC       In vitro, it binds MAPT via the MT-binding domains (By
CC       similarity). Associates with microtubules in the presence of ATP
CC       and in a kinesin-dependent manner (By similarity). Interacts,
CC       through a C-terminal domain, with GNAO1. Amyloid beta-42 binds
CC       CHRNA7 in hippocampal neurons (By similarity). Beta-amyloid
CC       associates with HADH2 (By similarity).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein. Cell surface
CC       protein that rapidly becomes internalized via clathrin-coated
CC       pits. During maturation, the immature APP (N-glycosylated in the

  Query Match             100.0%;  Score 62;  DB 1;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 0.00062;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        724 KKKQYTSIHHG 734

Search completed: October  4, 2004, 18:57:18
Job time : 20.087 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:37:37 ; Search time 52.1304 Seconds
(without alignments)
66.577 Million cell updates/sec

Title: US-10-048-209-4
Perfect score: 62
Sequence: 1 KKKQYTSIHHG 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62	100.0	49	6	O97917	O97917 bos taurus
2	62	100.0	113	13	Q8JH58	Q8jh58 chelydra se
3	62	100.0	218	11	Q8BPV5	Q8bpv5 mus musculu
4	62	100.0	384	11	Q8BPC7	Q8bpc7 mus musculu
5	62	100.0	534	13	O93296	O93296 gallus gall
6	62	100.0	569	13	Q9PVL1	Q9pvl1 gallus gall
7	62	100.0	695	13	Q9DGJ8	Q9dgj8 gallus gall
8	62	100.0	699	13	O57394	O57394 narke japon
9	62	100.0	751	13	Q9DGJ7	Q9dgj7 gallus gall
10	59	95.2	239	13	Q8UUI7	Q8uui7 brachydanio
11	59	95.2	357	13	Q8UUI8	Q8uui8 brachydanio
12	59	95.2	472	13	Q8UUS0	Q8uus0 brachydanio
13	59	95.2	612	13	Q9I9E7	Q9i9e7 brachydanio
14	59	95.2	678	13	Q7ZZT1	Q7zzt1 brachydanio
15	59	95.2	693	13	Q98SG0	Q98sg0 xenopus lae
16	59	95.2	694	13	Q8UUR9	Q8uur9 brachydanio
17	59	95.2	695	13	Q98SF9	Q98sf9 xenopus lae
18	59	95.2	695	13	Q7ZXQ0	Q7zxq0 xenopus lae
19	59	95.2	738	13	Q90W28	Q90w28 brachydanio
20	59	95.2	747	13	Q91963	Q91963 xenopus. ap
21	50	80.6	82	4	Q16020	Q16020 homo sapien
22	50	80.6	82	4	Q16014	Q16014 homo sapien
23	50	80.6	82	4	Q16019	Q16019 homo sapien
24	42	67.7	133	5	O17443	O17443 brevicoryne
25	41	66.1	464	10	Q9FEP0	Q9fep0 adonis pala
26	40	64.5	251	13	Q8AWR6	Q8awr6 brachydanio
27	40	64.5	398	16	Q7V4T7	Q7v4t7 prochloroco
28	40	64.5	452	10	Q9SZ00	Q9sz00 arabidopsis
29	40	64.5	466	10	Q94B35	Q94b35 arabidopsis
30	39	62.9	71	16	Q816L5	Q81615 bacillus ce
31	39	62.9	164	16	Q81P90	Q81p90 bacillus an
32	39	62.9	578	5	Q8ICM3	Q8icm3 plasmodium
33	38	61.3	228	16	Q8A0P0	Q8a0p0 bacteroides
34	38	61.3	289	16	Q8RIL2	Q8ril2 fusobacteri
35	38	61.3	451	2	Q9RCU3	Q9rcu3 staphylococ
36	38	61.3	451	16	Q8CTZ4	Q8ctz4 staphylococ
37	38	61.3	1495	5	Q8IKI3	Q8iki3 plasmodium
38	37	59.7	71	16	Q83DX8	Q83dx8 coxiella bu
39	37	59.7	158	10	Q8GWR9	Q8gwr9 arabidopsis
40	37	59.7	208	11	Q8R0R7	Q8r0r7 mus musculu
41	37	59.7	210	10	Q9LSU4	Q9lsu4 arabidopsis
42	37	59.7	264	13	Q804V7	Q804v7 brachydanio
43	37	59.7	267	10	Q39092	Q39092 arabidopsis
44	37	59.7	271	13	Q8AW31	Q8aw31 brachydanio
45	37	59.7	272	13	Q8JFS9	Q8jfs9 brachydanio

ALIGNMENTS

RESULT 1
O97917
ID O97917 PRELIMINARY; PRT; 49 AA.
AC O97917;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```
DE   Amyloid protein (Fragment).
GN   APP.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20063685; PubMed=10594237;
RA   Konfortov B.A., Licence V.E., Miller J.R.;
RT   "Resequencing of DNA from a diverse panel of cattle reveals a high
RT   level of polymorphism in both intron and exon.";
RL   Mamm. Genome 10:1142-1145(1999).
DR   EMBL; AJ133033; CAB38017.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
FT   NON_TER       1      1
FT   NON_TER      49     49
SQ   SEQUENCE   49 AA;  5183 MW;  6287463F0559BDED CRC64;

  Query Match             100.0%;  Score 62;  DB 6;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 0.00039;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db         36 KKKQYTSIHHG 46


RESULT 2
Q8JH58
ID   Q8JH58      PRELIMINARY;      PRT;   113 AA.
AC   Q8JH58;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid beta protein (Fragment).
OS   Chelydra serpentina serpentina (common snapping turtle).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Testudines; Cryptodira; Testudinoidea; Chelydridae; Chelydra.
OX   NCBI_TaxID=134619;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21876906; PubMed=11882478;
RA   Trudeau V.L., Chiu S., Kennedy S.W., Brooks R.J.;
RT   "Octylphenol (OP) alters the expression of members of the amyloid
RT   protein family in the hypothalamus of the snapping turtle, Chelydra
RT   serpentina serpentina.";
RL   Environ. Health Perspect. 110:269-275(2002).
DR   EMBL; AF541917; AAN04908.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   113 AA;  12750 MW;  72515C930496E053 CRC64;

  Query Match             100.0%;  Score 62;  DB 13;  Length 113;
  Best Local Similarity   100.0%;  Pred. No. 0.00091;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db         67 KKKQYTSIHHG 77


RESULT 3
Q8BPV5
ID   Q8BPV5      PRELIMINARY;      PRT;   218 AA.
AC   Q8BPV5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid beta (Fragment).
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Lung;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK052448; BAC34997.1; -.
DR   MGD; MGI:88059; App.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1   ? 1
SQ   SEQUENCE   218 AA;  24118 MW;  95B55AFDAE1D0EF5 CRC64;

  Query Match             100.0%;  Score 62;  DB 11;  Length 218;
  Best Local Similarity   100.0%;  Pred. No. 0.0018;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        172 KKKQYTSIHHG 182


RESULT 4
```

```
Q8BPC7
ID   Q8BPC7      PRELIMINARY;      PRT;   384 AA.
AC   Q8BPC7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amyloid beta (Fragment).
GN   APP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK076506; BAC36369.1; -.
DR   MGD; MGI:88059; App.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   384 AA;  43990 MW;  A81B1AD8AE683173 CRC64;

  Query Match             100.0%;  Score 62;  DB 11;  Length 384;
  Best Local Similarity   100.0%;  Pred. No. 0.0031;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        338 KKKQYTSIHHG 348


RESULT 5
O93296
ID   O93296      PRELIMINARY;      PRT;   534 AA.
AC   O93296;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98337885; PubMed=9671674;
RA   Barnes N.Y., Li L., Yoshikawa K., Schwartz L.M., Oppenheim R.W.,
RA   Milligan C.E.;
RT   "Increased production of amyloid precursor protein provides a
RT   substrate for caspase-3 in dying motoneurons.";
RL   J. Neurosci. 18:5869-5880(1998).
DR   EMBL; AF042098; AAC25052.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   534 AA;  60597 MW;  FB53ECC2E66D4C92 CRC64;

  Query Match             100.0%;  Score 62;  DB 13;  Length 534;
  Best Local Similarity   100.0%;  Pred. No. 0.0044;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        488 KKKQYTSIHHG 498


RESULT 6
Q9PVL1
ID   Q9PVL1      PRELIMINARY;      PRT;   569 AA.
AC   Q9PVL1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Amyloid protein (Fragment).
GN   APP.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Coulson E.J., Paliga K., Beyreuther K., Masters C.L.;
RT   "What the evolution of the amyloid protein precursor supergene family
RT   tells us about its function.";
RL   Neurochem. Int. 0:0-0(2000).
DR   EMBL; AF030341; AAF12698.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
```

```
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   569 AA;  64753 MW;  0AB8BB851863A19D CRC64;

  Query Match             100.0%;  Score 62;  DB 13;  Length 569;
  Best Local Similarity   100.0%;  Pred. No. 0.0046;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        524 KKKQYTSIHHG 534


RESULT 7
Q9DGJ8
ID   Q9DGJ8      PRELIMINARY;      PRT;   695 AA.
AC   Q9DGJ8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Beta-amyloid precursor protein 695 isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sarasa M., Rodolosse A., Sorribas V.;
RT   "Cloning of full-length chicken beta-amyloid precursor protein
RT   isoforms.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289218; AAG00593.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   695 AA;  78565 MW;  F201ED02AEC86D95 CRC64;

  Query Match             100.0%;  Score 62;  DB 13;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. 0.0057;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        649 KKKQYTSIHHG 659


RESULT 8
O57394
ID   O57394      PRELIMINARY;      PRT;   699 AA.
AC   O57394;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   EL amyloid precursor protein 699.
GN   EL APP699.
OS   Narke japonica (Electric ray).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
OC   Torpediniformes; Narcinoidei; Narkidae; Narke.
OX   NCBI_TaxID=62965;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Electric lobe;
RX   MEDLINE=98129705; PubMed=9461486;
RA   Iijima K., Lee D.-S., Okutsu J., Tomita S., Hirashima N., Kirino Y.,
RA   Suzuki T.;
RT   "cDNA isolation of Alzheimer's amyloid precursor protein from
RT   cholinergic nerve terminals of the electric organ of the electric
RT   ray.";
RL   Biochem. J. 330:29-33(1998).
DR   EMBL; AB005544; BAA24230.1; -.
DR   HSSP; P05067; 1HZ3.
DR   GO; GO:0016020; C:membrane; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
SQ   SEQUENCE   699 AA;  78879 MW;  952915C309D50E5C CRC64;

  Query Match             100.0%;  Score 62;  DB 13;  Length 699;
  Best Local Similarity   100.0%;  Pred. No. 0.0057;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        653 KKKQYTSIHHG 663


RESULT 9
Q9DGJ7
ID   Q9DGJ7      PRELIMINARY;      PRT;   751 AA.
AC   Q9DGJ7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Beta-amyloid precursor protein 751 isoform.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sarasa M., Rodolosse A., Sorribas V.;
RT   "Cloning of full-length chicken beta-amyloid precursor protein
RT   isoforms.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF289219; AAG00594.1; -.
DR   HSSP; P05067; 1BA4.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004867; F:serine protease inhibitor activity; IEA.
DR   InterPro; IPR008155; A4_APP.
DR   InterPro; IPR008154; A4_extra.
DR   InterPro; IPR001255; Beta-APP.
DR   InterPro; IPR002223; Kunitz_BPTI.
DR   Pfam; PF02177; A4_EXTRA; 1.
DR   Pfam; PF03494; Beta-APP; 1.
DR   Pfam; PF00014; Kunitz_BPTI; 1.
DR   PRINTS; PR00203; AMYLOIDA4.
DR   PRINTS; PR00759; BASICPTASE.
DR   ProDom; PD000222; Kunitz_BPTI; 1.
DR   SMART; SM00006; A4_EXTRA; 1.
DR   SMART; SM00131; KU; 1.
DR   PROSITE; PS00319; A4_EXTRA; 1.
DR   PROSITE; PS00320; A4_INTRA; 1.
DR   PROSITE; PS00280; BPTI_KUNITZ_1; 1.
DR   PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW   Protease inhibitor; Serine protease inhibitor.
SQ   SEQUENCE   751 AA;  84705 MW;  E78E9413A8033D84 CRC64;

  Query Match            100.0%;  Score 62;  DB 13;  Length 751;
  Best Local Similarity  100.0%;  Pred. No. 0.0061;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KKKQYTSIHHG 11
              ||||||||||| 
Db        705 KKKQYTSIHHG 715

Search completed: October  4, 2004, 18:56:24
Job time : 54.1304 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:37:13 ; Search time 92.8696 Seconds
(without alignments)
48.679 Million cell updates/sec

Title: US-10-048-209-5
Perfect score: 92
Sequence: 1 KQIKIWFQNRRMKWKK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	92	100.0	16	4	AAB67795	Aab67795 Amino aci
2	89	96.7	16	2	AAW45974	Aaw45974 Cysteine
3	89	96.7	16	2	AAW33407	Aaw33407 Peptide 4
4	89	96.7	16	2	AAW33410	Aaw33410 D-form pe
5	89	96.7	16	2	AAW82958	Aaw82958 Oestrogen
6	89	96.7	16	2	AAW56397	Aaw56397 Preferred
7	89	96.7	16	2	AAW71270	Aaw71270 Antennape
8	89	96.7	16	2	AAW71316	Aaw71316 Antennape
9	89	96.7	16	2	AAW30508	Aaw30508 Drosophil
10	89	96.7	16	2	AAW91046	Aaw91046 Internali
11	89	96.7	16	2	AAY52102	Aay52102 Peptide f
12	89	96.7	16	2	AAY00859	Aay00859 Peptide p
13	89	96.7	16	2	AAY13509	Aay13509 Signal se
14	89	96.7	16	3	AAY87920	Aay87920 Drosophil
15	89	96.7	16	3	AAB27060	Aab27060 Beta-cate
16	89	96.7	16	3	AAY93667	Aay93667 Peptide w
17	89	96.7	16	3	AAY67966	Aay67966 Carboxyfl
18	89	96.7	16	3	AAY93551	Aay93551 Amino aci
19	89	96.7	16	3	AAY55818	Aay55818 Signal se
20	89	96.7	16	3	AAY71008	Aay71008 Drosophil
21	89	96.7	16	3	AAY51212	Aay51212 Antennape
22	89	96.7	16	3	AAY51167	Aay51167 Drosophil
23	89	96.7	16	3	AAB10343	Aab10343 Peptide A
24	89	96.7	16	3	AAB19251	Aab19251 Fragment
25	89	96.7	16	3	AAY93178	Aay93178 Protegrin
26	89	96.7	16	3	AAB35694	Aab35694 Peptide a
27	89	96.7	16	3	AAB22025	Aab22025 Membrane
28	89	96.7	16	3	AAB29423	Aab29423 ANTP pept
29	89	96.7	16	3	AAB03927	Aab03927 Internali
30	89	96.7	16	3	AAY93954	Aay93954 Peptide u
31	89	96.7	16	3	AAB29574	Aab29574 Antennape
32	89	96.7	16	3	ADE14785	Ade14785 Carrier m
33	89	96.7	16	3	ADE14761	Ade14761 Drosophil
34	89	96.7	16	4	AAB73091	Aab73091 Rheumatoi
35	89	96.7	16	4	AAB60004	Aab60004 Internali
36	89	96.7	16	4	AAB70753	Aab70753 Cell memb
37	89	96.7	16	4	AAE02974	Aae02974 Protein t
38	89	96.7	16	4	AAB60671	Aab60671 Antennape
39	89	96.7	16	4	AAU06064	Aau06064 Drosophil
40	89	96.7	16	4	AAB49914	Aab49914 HIF-1alph
41	89	96.7	16	4	AAB66996	Aab66996 Antennape
42	89	96.7	16	4	AAU00813	Aau00813 Fruit fly
43	89	96.7	16	4	AAE12205	Aae12205 Membrane
44	89	96.7	16	5	ABB78030	Abb78030 Peptide d
45	89	96.7	16	5	ABG78985	Abg78985 Cell pene

ALIGNMENTS

RESULT 1
AAB67795
ID AAB67795 standard; peptide; 16 AA.
XX
AC AAB67795;
XX
DT 11-JUN-2001 (first entry)
XX
DE Amino acid sequence of an internalisation peptide.
XX
KW Amyloid protein precursor; APP; apoptosis; cancer; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200109170-A1.
XX

PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-FR002174.
XX
PR 30-JUL-1999; 99FR-00009929.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Allinquant B, Prochiantz A;
XX
DR WPI; 2001-257398/26.
XX
PT Peptides derived from the cytoplasmic domain of the amyloid protein
PT precursor, useful in the treatment of cancer and Alzheimer's disease.
XX
PS Claim 3; Page 13; 28pp; French.
XX
CC The present sequence represents an internalisation peptide. It may be
CC attached to peptides derived from the cytoplasmic domain of the human
CC amyloid protein precursor (APP). APP peptides derived from the
CC cytoplasmic domain, and containing the membrane domain juxtaposed to the
CC cytoplasmic domain of APP are useful for selecting and screening products
CC capable of inhibiting apoptosis. The peptides are useful in the treatment
CC of cancer and Alzheimer's disease
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 92; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
||||||||||||||||
Db 1 KQIKIWFQNRRMKWKK 16

RESULT 2
AAW45974
ID AAW45974 standard; peptide; 16 AA.
XX
AC AAW45974;
XX
DT 01-JUL-1998 (first entry)
XX
DE Cysteine protease inhibiting peptide for preventing cell death.
XX
KW Neuronal cell death; neurodegenerative disorder; inhibition;
KW cysteine protease; cardiovascular; liver disease.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-3-nitro-2-pyridyl-sulphenyl-Arg"
XX
PN WO9735876-A1.
XX

PD 02-OCT-1997.
XX
PF 04-MAR-1997; 97WO-US004158.
XX
PR 04-MAR-1996; 96US-00610220.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Troy CM;
XX
DR WPI; 1997-489561/45.
XX
PT New cysteine protease inhibiting peptide(s) for preventing cell death -
PT in cases of neuro:degenerative, cardiovascular and liver diseases, and
PT their peptido:mimetics, and general method for identifying enzyme
PT inhibiting peptides.
XX
PS Claim 8; Page 68; 112pp; English.
XX
CC This sequence represents a specifically claimed peptide of the formula: V
CC -(AA1)n-Cys(V'')-(AA2)m-V' (I), in which n and m = 0-5, totalling 2-5; if
CC n =1, AA1 = Ala; if n = 2, (AA1)n = Gln-Ala; and if n =3 or more, (AA1)n
CC = (X)p-Gln-Ala; X = any amino acid; p = 1-3, depending on value of n; if
CC m = 1, AA2 = Arg; if m =2, (AA2)n = Arg-Gly; if m = 3 or more, (AA2)n =
CC Arg-Gly-(X)q; q = 1-3, depending on value of m; V, V' and V'', any or all
CC of which may be absent, = agent able to direct the compound to a specific
CC cell. The peptides are inhibitors of cysteine proteases, specifically
CC interleukin-1 beta converting enzyme (ICE). They inhibit death of cells,
CC particularly in humans, and can be used to treat neurodegenerative
CC diseases (e.g. ageing, Alzheimer's, Machado-Joseph, Parkinson's or
CC Huntington's diseases, multiple sclerosis, muscular dystrophy, stroke),
CC cardiovascular disease and liver disorders. The peptides should be more
CC specific than pseudosubstrate inhibitors
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:|||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 3
AAW33407
ID AAW33407 standard; peptide; 16 AA.
XX
AC AAW33407;
XX
DT 27-AUG-2003 (revised)
DT 17-MAR-1998 (first entry)
XX
DE Peptide 43-58 of homeodomain Antp.
XX
KW homeodomain; transcription factor; Antennapedia; Antp; vector;

KW transfection; hydrophobic.
XX
OS Unidentified.
XX
PN WO9712912-A1.
XX
PD 10-APR-1997.
XX
PF 04-OCT-1996; 96WO-FR001553.
XX
PR 05-OCT-1995; 95FR-00011714.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Chassaing G, Prochiantz A;
XX
DR WPI; 1997-226166/20.
XX
PT New peptide(s) of high hydrophobic amino acid content - useful as vectors
PT for delivering peptides and nucleic acids to cells.
XX
PS Claim 1; Page 7; 35pp; French.
XX
CC New peptides are provided which are 16 amino acids long and which are
CC analogues of the peptide corresponding to residues 43-58 of the
CC Antennapedia transcription factor homeodomain (AntpHD). The peptides
CC contain 6-10 hydrophobic amino acids. They have the general formula: X1-
CC X2-X3-X4-X5-Trp-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16 or X16-X15-X14-X13-
CC X12-X11-X10-X9-X8-X7-Trp-X5-X4-X3-X2-X1 in which X1-X5 and X7-X16 are any
CC alpha-amino acids, provided that: (1) the peptide contains 6-10
CC hydrophobic amino acids; (2) X3 and X5 are not both Val; and (3) the
CC natural Antp 43-58 sequence RQIKIWFQNRRMKWKK (i.e. the present sequence)
CC is excluded. Specific examples of these peptides are given in AAW33408 -
CC AAW33416. The peptides are used as vectors for introducing into live
CC cells compounds which affect cell function, esp. peptides and nucleic
CC acids. They can cross cellular membranes and reach various cell
CC compartments. They are as effective as helix 3 of a homeodomain peptide.
CC (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db          1 RQIKIWFQNRRMKWKK 16
```

RESULT 4
AAW33410
ID AAW33410 standard; peptide; 16 AA.
XX
AC AAW33410;
XX
DT 17-MAR-1998 (first entry)
XX
DE D-form peptide 43-58 of homeodomain Antp.
XX
KW homeodomain; transcription factor; Antennapedia; Antp; vector;
KW transfection; hydrophobic.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 1. .16
FT /note= "all residues are D-form"
FT Modified-site 1
FT /note= "in determining the ability of this sequence to be
FT internalised in cells, a biotin-aminopentanoyl group was
FT attached to the N-terminal"
XX
PN WO9712912-A1.
XX
PD 10-APR-1997.
XX
PF 04-OCT-1996; 96WO-FR001553.
XX
PR 05-OCT-1995; 95FR-00011714.
XX
PA (CNRS) CNRS CENT NAT RECH SCI.
XX
PI Chassaing G, Prochiantz A;
XX
DR WPI; 1997-226166/20.
XX
PT New peptide(s) of high hydrophobic amino acid content - useful as vectors
PT for delivering peptides and nucleic acids to cells.
XX
PS Disclosure; Page 7; 35pp; French.
XX
CC New peptides are provided which are 16 amino acids long and which are
CC analogues of the peptide corresponding to residues 43-58 of the
CC Antennapedia transcription factor homeodomain (AntpHD). The peptides
CC contain 6-10 hydrophobic amino acids. They have the general formula: X1-
CC X2-X3-X4-X5-Trp-X7-X8-X9-X10-X11-X12-X13-X14-X15-X16 or X16-X15-X14-X13-
CC X12-X11-X10-X9-X8-X7-Trp-X5-X4-X3-X2-X1 in which X1-X5 and X7-X16 are any
CC alpha-amino acids, provided that: (1) the peptide contains 6-10
CC hydrophobic amino acids; (2) X3 and X5 are not both Val; and (3) the
CC natural Antp 43-58 sequence RQIKIWFQNRRMKWKK (see AAW33407) is excluded.
CC The present sequence (the D-form of the 43-58 peptide) is a specific
CC example of the new peptides. The peptides are used as vectors for
CC introducing into live cells compounds which affect cell function,
CC especially peptides and nucleic acids. They can cross cellular membranes
CC and reach various cell compartments. They are as effective as helix 3 of
CC a homeodomain peptide
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db          1 RQIKIWFQNRRMKWKK 16
```

RESULT 5
AAW82958
ID AAW82958 standard; peptide; 16 AA.
XX
AC AAW82958;
XX
DT 04-FEB-1999 (first entry)
XX
DE Oestrogen receptor activity inhibiting peptide #14.
XX
KW Human; oestrogen receptor activity inhibitor; anti-oestrogen; diagnosis;
KW breast cancer; estrogen; tumour; phosphotyrosyl peptide;
KW malonyltyrosyl peptide; steroid receptor co-activator-1.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9846250-A1.
XX
PD 22-OCT-1998.
XX
PF 14-APR-1998; 98WO-US007711.
XX
PR 14-APR-1997; 97US-0043545P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Pietras RJ;
XX
DR WPI; 1998-594522/50.
XX
PT New anti-oestrogen peptide compositions - comprise sequences based on
PT oestrogen receptor and steroid receptor co-activator-1 sequences, used
PT for treating cancers.
XX
PS Claim 59; Page 156; 182pp; English.
XX
CC The present invention describes a composition comprising an isolated
CC oestrogen receptor activity inhibiting (anti-oestrogen) peptide. The
CC peptides used in the composition comprise sequences of human oestrogen
CC receptor (OR) surrounding Tyr537 and steroid receptor co-activator-1 (SCR
CC -1). The peptide compositions, nucleic acids and vectors of the present
CC invention can reduce OR activity in a cell, reduced OR polypeptide
CC dimerisation in a cell and reduce the binding of SRC-1 polypeptide to an
CC OR polypeptide dimer in a cell. They can be used for killing cancer cells
CC and treating cancers, particularly breast cancer. The present sequence
CC represents a specifically claimed anti-oestrogen peptide
XX
SQ Sequence 16 AA;

  Query Match   96.7%; Score 89; DB 2; Length 16;
  Best Local Similarity   93.8%; Pred. No. 2.7e-06;
  Matches   15; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db          1 RQIKIWFQNRRMKWKK 16
```

RESULT 6
AAW56397
ID AAW56397 standard; peptide; 16 AA.
XX
AC AAW56397;
XX
DT 05-AUG-1998 (first entry)
XX
DE Preferred signal sequence of the invention.
XX
KW Signal peptide; nuclear localisation signal; NLS;
KW immunosuppressive activity; inhibition; nuclear translocation inhibitor;
KW nuclear translocation; treatment; immune disorder; autoimmune disease;
KW hypersensitivity; sepsis; prevention; septic shock; antiviral agent;
KW tumour growth suppressor.
XX
OS Unidentified.
XX
PN WO9811907-A1.
XX
PD 26-MAR-1998.
XX
PF 15-SEP-1997; 97WO-US016217.
XX
PR 20-SEP-1996; 96US-0026978P.
PR 12-SEP-1997; 97US-00928958.
XX
PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
PI Nadler SG, Cleaveland JS, Blake J, Haffar OK;
XX
DR WPI; 1998-217028/19.
XX
PT Nuclear translocation inhibitor polypeptides - comprising signal sequence
PT for delivery through the cytoplasmic membrane and at least 2 nuclear
PT localisation sequences.
XX
PS Claim 5; Page 43; 69pp; English.
XX
CC Peptides AAW56397-99 represent preferred signal sequences of the
CC invention. They are used to construct the nuclear translocation inhibitor
CC polypeptides of the invention. Nuclear translocation inhibitor
CC polypeptides comprise a signal sequence peptide capable of delivering the
CC polypeptide through the cytoplasmic membrane into a cell, and at least 2
CC nuclear localisation signals (NLSs). The polypeptides can be used to
CC inhibit nuclear translocation of a cellular protein. In addition, since
CC the nuclear translocation of certain cellular peptides is required for
CC the host organism to mount an immune response, the polypeptide inhibitors
CC are useful as immunosuppression agents. The polypeptides can therefore be
CC used for the treatment of immune disorders including autoimmune diseases.

CC The polypeptides can also be used for treating physical symptoms
CC manifested by responses to allergens which can initiate a state of
CC hypersensitivity, for the treatment of sepsis and in the prevention of
CC septic shock, antiviral agents, tumour growth suppressors, and for
CC transcriptionally modulating the expression of cellular genes
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:||||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 7
AAW71270
ID AAW71270 standard; protein; 16 AA.
XX
AC AAW71270;
XX
DT 23-NOV-1998 (first entry)
XX
DE Antennapedia peptide for directing antisense oligonuclectides to a cell.
XX
KW Antisense oligonucleotide; inhibit; cell death; Nedd2; human Ich-1L gene;
KW neuronal cell death; treatment; aging; amyotrophic lateral sclerosis;
KW Alzheimer's disease; dentatorubral; pallidolyusial atrophy;
KW Huntington's disease; Machoado-Joseph disease; multiple sclerosis;
KW muscular dystrophy; Parkinson's disease; senility;
KW spinocerebellar ataxia type I; spinobulbar muscular atrophy; stroke;
KW trauma; antennapedia.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal protected with 3-nitro-2-pyridyl
FT sulphenyl group (NpyS)"
XX
PN WO9838861-A1.
XX
PD 11-SEP-1998.
XX
PF 03-MAR-1998; 98WO-US004128.
XX
PR 03-MAR-1997; 97US-00810540.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Troy CM, Shelanski ML;
XX
DR WPI; 1998-506333/43.
XX
PT Anti-sense oligonucleotide(s) to cysteine aspartase genes - used to
PT inhibit neurodegenerative disorder associated with e.g. ageing,
PT Alzheimer's, Huntington's or Parkinson's disease.
XX
PS Claim 5; Page 39; 60pp; English.
XX
CC AAW71270 and AAW71315-16 represent antennapedia peptides which are used
CC to, direct the antisense oligonucleotides (AAV54973-74) of the invention
CC to a cell. The antisense oligonucleotides are used to inhibit cell death
CC mediated by withdrawal of a trophic factor. AAV54973 inhibits the
CC expression of a Nedd2 rodent gene, while AAV54974 inhibits expression of
CC a human Ich-1L gene. The oligonucleotides are used to inhibit neuronal
CC cell death, especially for treatment of neuronal cell death caused by
CC e.g. aging, amyotrophic lateral sclerosis, Alzheimer's disease,
CC dentatorubral and pallidolyusial atrophy, Huntington's disease, Machoado-
CC Joseph disease, multiple sclerosis, muscular dystrophy, Parkinson's
CC disease, senility, spinocerebellar ataxia type I, spinobulbar muscular
CC atrophy, stroke or trauma
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:||||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 8
AAW71316
ID AAW71316 standard; protein; 16 AA.
XX
AC AAW71316;
XX
DT 23-NOV-1998 (first entry)
XX
DE Antennapedia peptide for directing antisense oligonucleotides to a cell.
XX
KW Antisense oligonucleotide; inhibit; cell death; Nedd2; human Ich-1L gene;
KW neuronal cell death; treatment; aging; amyotrophic lateral sclerosis;
KW Alzheimer's disease; dentatorubral; pallidolyusial atrophy;
KW Huntington's disease; Machoado-Joseph disease; multiple sclerosis;
KW muscular dystrophy; Parkinson's disease; senility;
KW spinocerebellar ataxia type I; spinobulbar muscular atrophy; stroke;
KW trauma; antennapedia.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal protected with 3-nitro-2-pyridyl
FT sulphenyl group (NpyS)"
XX
PN WO9838861-A1.
XX
PD 11-SEP-1998.

XX
PF 03-MAR-1998; 98WO-US004128.
XX
PR 03-MAR-1997; 97US-00810540.
XX
PA (UYCO) UNIV COLUMBIA NEW YORK.
XX
PI Troy CM, Shelanski ML;
XX
DR WPI; 1998-506333/43.
XX
PT Anti-sense oligonucleotide(s) to cysteine aspartase genes - used to
PT inhibit neurodegenerative disorder associated with e.g. ageing,
PT Alzheimer's, Huntington's or Parkinson's disease.
XX
PS Disclosure; Page 14; 60pp; English.
XX
CC AAW71270 and AAW71315-16 represent antennapedia peptides which are used
CC to, direct the antisense oligonucleotides (AAV54973-74) of the invention
CC to a cell. The antisense oligonucleotides are used to inhibit cell death
CC mediated by withdrawal of a trophic factor. AAV54973 inhibits the
CC expression of a Nedd2 rodent gene, while AAV54974 inhibits expression of
CC a human Ich-1L gene. The oligonucleotides are used to inhibit neuronal
CC cell death, especially for treatment of neuronal cell death caused by
CC e.g. aging, amyotrophic lateral sclerosis, Alzheimer's disease,
CC dentatorubral and pallidolysial atrophy, Huntington's disease, Machoado-
CC Joseph disease, multiple sclerosis, muscular dystrophy, Parkinson's
CC disease, senility, spinocerebellar ataxia type I, spinobulbar muscular
CC atrophy, stroke or trauma
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:||||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 9
AAW30508
ID AAW30508 standard; peptide; 16 AA.
XX
AC AAW30508;
XX
DT 26-OCT-1998 (first entry)
XX
DE Drosophila membrane translocation sequence.
XX
KW DP-1; transcription factor; antagonist; E2F protein; apoptosis;
KW cell proliferation; cardiovascular cell; restenosis; tumour;
KW surgical stent; therapy; membrane translocation; antennapedia protein.
XX
OS Drosophila melanogaster.
XX

PN WO9828334-A1.
XX
PD 02-JUL-1998.
XX
PF 22-DEC-1997; 97WO-GB003506.
XX
PR 20-DEC-1996; 96GB-00026589.
XX
PA (PROL-) PROLIFIX LTD.
XX
PI La Thangue NB, Bandara LR;
XX
DR WPI; 1998-377596/32.
XX
PT Polypeptide fragments of the DP-1 transcription factor - used for
PT inducing apoptosis, specifically in tumour and cardiovascular cells, e.g.
PT for preventing re-stenosis.
XX
PS Disclosure; Page 5; 55pp; English.
XX
CC This polypeptide comprises a membrane translocation sequence derived from
CC the Drosophila melanogaster antennapedia protein. Such membrane
CC translocation sequences are useful in directing entry of a polypeptide
CC into a cell. Polypeptides of the invention (see AAW30504-07) are derived
CC from the DEF box region (see AAW30501) of transcription factor DP1. They
CC act as antagonists of the heterodimerisation of a DP protein with an E2F
CC protein, and can be used to induced apoptosis, specifically in tumour and
CC cardiovascular cells, e.g. for preventing restenosis. A claimed fusion
CC protein comprises a DEF box peptide and the Drosophila antennapedia
CC protein membrane translocation sequence
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:||||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 10
AAW91046
ID AAW91046 standard; peptide; 16 AA.
XX
AC AAW91046;
XX
DT 24-MAR-1999 (first entry)
XX
DE Internalization sequence associated with cadherin modulating agents.
XX
KW Alpha-catenin; beta-catenin; interaction; modulation; cell adhesion;
KW cadherin-mediated function; demyelinating neurological disease;
KW multiple sclerosis; drug delivery; cancer; angiogenesis; immune system;
KW central nervous system; apoptosis induction; cadherin-expression cell;
KW pregnancy prevention; vasopermeability; synaptic stability; diabetes;

KW rheumatoid arthritic; allergic response; learning; antennapedia protein;
KW antibody-mediated graft rejection; internalization sequence; memory.
XX
OS Synthetic.
XX
PN WO9845319-A2.
XX
PD 15-OCT-1998.
XX
PF 14-APR-1998; 98WO-CA000322.
XX
PR 10-APR-1997; 97US-0043361P.
XX
PA (UYMC-) UNIV MCGILL.
XX
PI Blaschuk OW, Gour BJ;
XX
DR WPI; 1999-024009/02.
XX
PT New catenin modulating agents - comprising peptides having a sequence HAV
PT or analogues or antibodies, used for modulating cadherin-mediated
PT functions.
XX
PS Claim 16; Page 78; 106pp; English.
XX
CC The present sequence represents antennapedia protein derived
CC internalization sequence associated with cadherin modulating agents.
CC These agents are used for modulating cadherin-mediated functions. They
CC can be used for disrupting interaction between alpha-catenin and beta-
CC catenin in a cell, inhibiting cell adhesion, e.g. between epithelial
CC cells, endothelial cells, neural cells, tumour cells and lymphocytes, for
CC treating a demyelinating neurological disease, e.g. multiple sclerosis,
CC for reducing unwanted cellular adhesion in a mammal, for enhancing the
CC delivery of a drug through the skin of a mammal, for enhancing the
CC delivery of a drug to a tumour in a mammal, for treating cancer in a
CC mammal, for inhibiting angiogenesis in a mammal, for enhancing drug
CC delivery to the central nervous system of a mammal, for inducing
CC apoptosis in a cadherin-expression cell, for modulating the immune system
CC of a mammal, for preventing pregnancy in a mammal, for increasing
CC vasopermeability in a mammal, or for inhibiting synaptic stability in a
CC mammal. In particular they can be used for treating diabetes, rheumatoid
CC arthritic, allergic responses, antibody-mediated graft rejection or for
CC stimulating learning and memory
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:||||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 11
AAY52102
ID AAY52102 standard; peptide; 16 AA.
XX
AC AAY52102;
XX
DT 28-JAN-2000 (first entry)
XX
DE Peptide from the third helix of antennapedia homeodomain protein.
XX
KW Rho binding region; Rho binding protein; ROCK II; ROCK I; Kinectin;
KW ROCK/Kinectin homology domain; cell growth; assay; Skn 7; modulator;
KW antifungal agent; antennapedia homeodomain protein; translocate.
XX
OS Drosophila melanogaster.
XX
PN WO9952941-A2.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1999; 99WO-GB001096.
XX
PR 09-APR-1998; 98GB-00007848.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX
PI Johnston LH, Treisman RH;
XX
DR WPI; 1999-620374/53.
XX
PT Assay for modulators of binding of Rho protein with its binding partners,
PT potentially useful as antifungal agents.
XX
PS Disclosure; Page 16; 68pp; English.
XX
CC This sequence is a peptide from the third helix of the Drosophila
CC antennapedia protein that translocates across biological membranes. This
CC sequence can be linked to the Rho binding region of the invention
CC (AAY52100) to cause translocation across eukaryotic cell membranes. The
CC Rho binding region is present in many Rho binding proteins, such as ROCK-
CC I, ROCK-II and Kinectin and is also referred to as the ROCK/Kinectin
CC homology domain. Rho-A is involved in many cellular processes including
CC stress fibre formation, cell motility, cytokinesis and apoptosis. The Rho
CC binding domain is used in the invention which relates to an assay for
CC potential modulators of cell growth. The assay consists of a Skn7
CC polypeptide, a beta2 polypeptide or a polypeptide consisting of a
CC ROCK/kinectin homology domain, a Rho polypeptide which binds to the
CC previous peptide and a test compound. The modulation of the binding of
CC the two peptides caused by the test compound is measured. The method is
CC used to identify modulators, particularly antagonists, of the interaction
CC of Rho with its binding partners, e.g. to render pathogens sensitive to
CC the host's defence system, particularly as antifungal agents. The
CC ROCK/kinectin homology domain can be used to screen databases for other
CC Rho-binding proteins
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;

```
  Best Local Similarity   93.8%;  Pred. No. 2.7e-06;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db          1 RQIKIWFQNRRMKWKK 16


RESULT 12
AAY00859
ID   AAY00859 standard; peptide; 16 AA.
XX
AC   AAY00859;
XX
DT   20-MAY-1999  (first entry)
XX
DE   Peptide pAntp(43-58) used in membrane-permeable construct.
XX
KW   Membrane-permeable construct; lipid membrane; membrane transport;
KW   oligonucleotide delivery; cancer therapy; signal transduction; inhibitor;
KW   gene therapy; transcription; translation; expression; replication.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   16
FT                   /note= "amidated"
XX
PN   WO9905302-A1.
XX
PD   04-FEB-1999.
XX
PF   16-JUL-1998;   98WO-US014761.
XX
PR   24-JUL-1997;   97US-0053678P.
XX
PA   (PEKE ) PERKIN-ELMER CORP.
XX
PI   Langel U,  Bartfai T,  Pooga M,  Valkna A,  Saar K,  Hallbrink M;
XX
DR   WPI; 1999-142952/12.
XX
PT   New membrane-permeable constructs - comprise a peptide linked by a labile
PT   bond to a nucleic acid analogue capable of hybridising with an
PT   intracellular polynucleotide.
XX
PS   Disclosure; Page 26; 60pp; English.
XX
CC   This sequence represents a peptide used in the construct of the
CC   invention. The construct is a membrane-permeable construct for transport
CC   across a lipid membrane, which comprises: (a) a nucleic acid analogue
CC   capable of hybridising with an intracellular polynucleotide (PN); (b) a
CC   peptide; and (c) a labile bond linking the nucleic acid analogue and the
CC   peptide. The membrane-permeable constructs can be used for delivery of
CC   oligonucleotides, nucleic acids and nucleic acid analogues into cells.
CC   They can be used for e.g. cancer therapy, signal transduction studies,
CC   identifying new intracellular drug targets or gene therapy. They can also
CC   be used for selectively inhibiting DNA transcription, RNA translation,
CC   RNA or DNA expression, DNA replication, or an DNA or RNA regulatory
CC   function of preselected DNA or RNA sequences in a living cell
XX
SQ   Sequence 16 AA;

  Query Match             96.7%;  Score 89;  DB 2;  Length 16;
  Best Local Similarity   93.8%;  Pred. No. 2.7e-06;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db          1 RQIKIWFQNRRMKWKK 16


RESULT 13
AAY13509
ID   AAY13509 standard; peptide; 16 AA.
XX
AC   AAY13509;
XX
DT   30-JUL-1999  (first entry)
XX
DE   Signal sequence of antennapedia.
XX
KW   Fusion protein; calpastatin; calpain; platelet aggregation; arthritis;
KW   hypoxia; erythrocyte sickling; sickle cell; HIV provirus; NF-kappaB;
KW   inflammation; asthma; immune response; restenosis; myocardial infarction;
KW   cancer; calpastat; antennapedia.
XX
OS   Mammalia.
XX
PN   WO9922756-A1.
XX
PD   14-MAY-1999.
XX
PF   04-NOV-1998;   98WO-US023526.
XX
PR   04-NOV-1997;   97US-00964302.
XX
PA   (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX
PI   Potter DA,  Skolnik PR;
XX
DR   WPI; 1999-326923/27.
XX
PT   Fusion protein of signal sequence and calpastatin.
XX
PS   Disclosure; Page 5; 46pp; English.
XX
CC   The invention features fusion proteins that contain a calpastatin peptide
CC   and a signal sequence capable of delivering the fusion protein into a
CC   eukaryotic cell. The fusion protein is used for the inhibition of calpain
CC   in a cell. The fusion proteins are specifically used: (a) to prevent
CC   aggregation and degranulation of platelets (e.g. during storage); (b) to
CC   inhibit hypoxia-induced sickling of erythrocytes (during storage,
CC   facilitating subsequent transfusion of autologous cells for treatment of
```

CC sickle cell crises); and (c) to inhibit activation of human immune
CC deficiency virus provirus in infected cells (or similarly for other
CC viruses regulated by NF-kappaB). Other disclosed uses are: to treat or
CC prevent inflammation (e.g. arthritis or asthma), unwanted immune
CC responses (e.g. transplant rejection), restenosis (associated with
CC angioplasty), cancer, subarachnoid hemorrhage, vasospasm, muscular
CC dystrophy, cataracts and traumatic birth injury; to prevent spread of
CC platelets on surfaces (e.g. when applied to the surface of stents,
CC catheters etc.); to reduce coronary thrombosis in by-pass surgery and
CC angioplasty; to treat myocardial infarction, or to prevent progression of
CC infarction (myocardial or cerebral). The fusion protein has a reversible
CC inhibitory effect and enters cells easily. It allows platelets to be
CC stored cold with reduced change in shape
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 2; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:||||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 14
AAY87920
ID AAY87920 standard; peptide; 16 AA.
XX
AC AAY87920;
XX
DT 11-SEP-2000 (first entry)
XX
DE Drosophila sp. antennapedia protein fragment.
XX
KW Rad51; cell proliferation; rad51TR1-131; rad51K-A134; immunosuppressive;
KW cytostatic; antiinflammatory; antiproliferative; programmed cell death;
KW treatment; autoimmune disorder; inflammation; cancer; graft rejection;
KW proliferative disorder; hyperproliferative disorder; antennapedia.
XX
OS Drosophila sp.
XX
PN US6057104-A.
XX
PD 02-MAY-2000.
XX
PF 05-NOV-1997; 97US-00964614.
XX
PR 05-NOV-1996; 96US-00758280.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Hasty P;
XX
DR WPI; 2000-349568/30.
XX
PT Mutant forms of mammalian Rad51 proteins and polypeptides that disrupt
PT cell proliferation and promote programmed cell death, encoded by specific
PT nucleotides.
XX
PS Disclosure; Col 31-32; 22pp; English.
XX
CC This invention describes novel truncated and altered forms of Rad51
CC products encoded by rad51TR1-131 and rad51K-A134. The products of the
CC invention have immunosuppressive, cytostatic, antiinflammatory and
CC antiproliferative activity. Altered and mutant forms of Rad51 are useful
CC for inhibiting cell proliferation and for promoting programmed cell
CC death. Therapeutic agents, factors or compounds capable of disrupting the
CC essential processes mediated by or associated with normal Rad51 or Rad52
CC are useful for treating diseases like autoimmune disorders, inflammation,
CC cancer, graft rejection and other proliferative and hyperproliferative
CC disorders. This sequence represents a fragment of the Drosophila
CC antennapedia protein which is described in the method of the invention
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 3; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
:||||||||||||||||
Db 1 RQIKIWFQNRRMKWKK 16

RESULT 15
AAB27060
ID AAB27060 standard; peptide; 16 AA.
XX
AC AAB27060;
XX
DT 15-FEB-2001 (first entry)
XX
DE Beta-catenin derived internalisation moiety SEQ ID NO: 47.
XX
KW Beta-catenin; cadherin-mediated intercellular adhesion;
KW cell differentiation; modulating agent; hair loss; skin exfoliation;
KW internalisation moiety; flanking sequence; transcription; hearing loss.
XX
OS Drosophila melanogaster.
XX
PN WO200053632-A1.
XX
PD 14-SEP-2000.
XX
PF 07-MAR-2000; 2000WO-CA000222.
XX
PR 09-MAR-1999; 99US-00265107.
XX
PA (UYMC-) UNIV MCGILL.
XX
PI Blaschuk OW, Gour BJ;
XX
DR WPI; 2000-594308/56.

XX
PT Stimulating beta-catenin mediated gene expression, cellular
PT differentiation and hair growth, involves contacting cells with
PT modulating agent capable of inhibiting interaction between alpha and beta
PT catenin.
XX
PS Claim 8; Page 38; 77pp; English.
XX
CC The present invention is concerned with methods of modulating the amount
CC of free beta-catenin in the cell, and methods of stimulating the
CC expression of genes involved in cellular differentiation, the
CC transcription of which is under the control of beta-catenin. The peptides
CC given in AAB27053-B27088, AAB27284-B27300 and AAB27330-B27351 can be used
CC as modulating agents which interrupt the interaction between alpha and
CC beta catenin, causing increased levels of the latter and stimulating the
CC activation of beta-catenin mediated transcription. This can be used to
CC stimulate cell differentiation, which can then be used to promote hair
CC growth and skin exfoliation. This latter is particularly useful in the
CC improvement of photodamaged skin and to minimise wrinkles. The modulating
CC peptide can also be used to reduce hearing loss resulting from inner ear
CC disorders such as hyperacusis and tinnitus
XX
SQ Sequence 16 AA;

Query Match 96.7%; Score 89; DB 3; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db          1 RQIKIWFQNRRMKWKK 16
```

Search completed: October 4, 2004, 18:52:48
Job time : 103.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:45:55 ; Search time 23.6522 Seconds
(without alignments)
65.071 Million cell updates/sec

Title: US-10-048-209-5
Perfect score: 92
Sequence: 1 KQIKIWFQNRRMKWKK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	89	96.7	33	2	S57235	antennapedia prote
2	89	96.7	42	2	I65241	homeotic protein H
3	89	96.7	45	2	PC1216	homeotic protein D
4	89	96.7	48	2	I51439	homeobox protein -
5	89	96.7	66	2	S15536	homeotic protein H
6	89	96.7	66	2	S15538	homeotic protein H
7	89	96.7	71	2	JC1161	homeotic protein 3
8	89	96.7	71	2	A60084	homeotic protein H
9	89	96.7	74	2	D34510	homeotic protein H
10	89	96.7	75	2	I51341	homeo box protein
11	89	96.7	75	2	S58852	homeotic protein S
12	89	96.7	76	2	C43559	homeotic protein R
13	89	96.7	78	2	I51342	homeo box protein

14	89	96.7	81	2	S47605	homeotic protein H
15	89	96.7	81	2	B29585	homeotic protein H
16	89	96.7	82	2	S08302	homeotic protein H
17	89	96.7	83	2	S47603	homeotic protein H
18	89	96.7	83	2	S50066	homeotic protein H
19	89	96.7	86	2	A34510	homeotic protein H
20	89	96.7	86	2	JT0489	homeotic protein Z
21	89	96.7	86	2	S08303	homeotic protein H
22	89	96.7	87	2	S00589	homeotic protein H
23	89	96.7	88	2	A03317	homeotic protein M
24	89	96.7	96	2	S08639	homeotic protein z
25	89	96.7	96	2	A05266	homeotic protein H
26	89	96.7	97	2	C27176	homeotic protein H
27	89	96.7	97	2	A24779	homeotic protein m
28	89	96.7	103	2	A32167	homeotic protein H
29	89	96.7	105	2	S47602	homeotic protein H
30	89	96.7	105	2	A27471	homeotic protein R
31	89	96.7	106	2	S36448	homeotic protein s
32	89	96.7	107	2	B61045	homeotic protein T
33	89	96.7	113	2	T10775	homeobox protein -
34	89	96.7	118	2	A24777	homeotic protein H
35	89	96.7	118	2	JT0273	homeotic protein H
36	89	96.7	118	2	B24777	homeotic protein M
37	89	96.7	119	2	A03314	homeotic protein m
38	89	96.7	138	2	S20087	homeotic protein b
39	89	96.7	148	2	PC4071	homeobox A5 protei
40	89	96.7	153	1	WJHU3C	homeotic protein H
41	89	96.7	153	1	WJMSX6	homeotic protein H
42	89	96.7	158	2	A27348	homeotic protein H
43	89	96.7	209	2	A43553	homeotic protein H
44	89	96.7	217	1	WJHU2C	homeotic protein H
45	89	96.7	217	1	WJMSX2	homeotic protein H

ALIGNMENTS

RESULT 1
S57235
antennapedia protein (clone p1105) - fruit fly (Drosophila pseudoobscura) (fragment)
C;Species: Drosophila pseudoobscura
C;Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 15-Oct-1999
C;Accession: S57235
R;Randazzo, F.M.; Seeger, M.A.; Huss, C.A.; Sweeney, M.A.; Cecil, J.K.; Kaufman, T.C.
Genetics 133, 319-330, 1993
A;Title: Structural changes in the antennapedia complex of Drosophila pseudoobscura.
A;Reference number: S57224
A;Accession: S57235
A;Molecule type: DNA
A;Residues: 1-33 <RAN>
A;Cross-references: EMBL:X77711
C;Genetics:
A;Gene: FlyBase:Antp
A;Cross-references: FlyBase:FBgn0012693
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-22/Domain: homeobox homology (fragment) <HOX>

Query Match 96.7%; Score 89; DB 2; Length 33;
Best Local Similarity 93.8%; Pred. No. 3.5e-07;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db          7 RQIKIWFQNRRMKWKK 22
```

RESULT 2
I65241
homeotic protein Hox-A - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 15-Oct-1999
C;Accession: I65241
R;Sakoyama, Y.; Mizuta, I.; Ogasawara, N.; Yoshikawa, H.
Biochem. Genet. 32, 351-360, 1994
A;Title: Cloning of rat homeobox genes.
A;Reference number: I52340; MUID:95217128; PMID:7702549
A;Accession: I65241
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-42 <RES>
A;Cross-references: GB:S76290; NID:g913077
C;Genetics:
A;Gene: Hox-A; Hox-1
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-40/Domain: homeobox homology (fragment) <HOX>

Query Match 96.7%; Score 89; DB 2; Length 42;
Best Local Similarity 93.8%; Pred. No. 4.4e-07;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         25 RQIKIWFQNRRMKWKK 40
```

RESULT 3
PC1216
homeotic protein DtHbx1 - planarian (Dugesia tigrina) (fragment)
C;Species: Dugesia tigrina
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Oct-1997
C;Accession: PC1216
R;Oliver, G.; Vispo, M.; Mailhos, A.; Martinez, C.; Sosa-Pineda, B.; Fielitz, W.; Ehrlich, R.
Gene 121, 337-342, 1992
A;Title: Homeoboxes in flatworms.
A;Reference number: JC1386; MUID:93077050; PMID:1359988
A;Accession: PC1216
A;Molecule type: DNA
A;Residues: 1-45 <OLI>

```
A;Cross-references: EMBL:X66822
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-45/Domain: homeobox homology (fragment) <HOX>

  Query Match             96.7%;  Score 89;  DB 2;  Length 45;
  Best Local Similarity   93.8%;  Pred. No. 4.7e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         30 RQIKIWFQNRRMKWKK 45


RESULT 4
I51439
homeobox protein - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 15-Oct-1999
C;Accession: I51439
R;Leroy, P.; DeRobertis, E.M.
Dev. Dyn. 194, 21-32, 1992
A;Title: Effects of lithium chloride and retinoic acid on the expression of
genes from the Xenopus laevis Hox 2 complex.
A;Reference number: I51439; MUID:93043517; PMID:1384809
A;Accession: I51439
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-48 <LER>
A;Cross-references: GB:M91587; NID:g214257; PIDN:AAA49750.1; PID:g214258
C;Genetics:
A;Gene: Hox2.2
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-27/Domain: homeobox homology (fragment) <HOX>

  Query Match             96.7%;  Score 89;  DB 2;  Length 48;
  Best Local Similarity   93.8%;  Pred. No. 5e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         12 RQIKIWFQNRRMKWKK 27


RESULT 5
S15536
homeotic protein Hox A7 - human (fragment)
N;Alternate names: homeotic protein Hox 1A
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 29-Aug-1997
C;Accession: S15536
R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino,
G.; Stornaiuolo, A.; Cafiero, M.; Faiella, A.; Simeone, A.
Genome 31, 745-756, 1989
A;Title: Organization of human class I homeobox genes.
A;Reference number: S15036; MUID:90215256; PMID:2576652
A;Accession: S15536
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-66 <BON>
C;Genetics:
A;Gene: GDB:HOXA7
A;Cross-references: GDB:120647; OMIM:142950
A;Map position: 7p15.3-7p15.3
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

  Query Match             96.7%;  Score 89;  DB 2;  Length 66;
  Best Local Similarity   93.8%;  Pred. No. 6.8e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         43 RQIKIWFQNRRMKWKK 58


RESULT 6
S15538
homeotic protein Hox A6 - human (fragment)
N;Alternate names: homeotic protein Hox 1B
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 29-Aug-1997
C;Accession: S15538
R;Boncinelli, E.; Acampora, D.; Pannese, M.; d'Esposito, M.; Somma, R.; Gaudino,
G.; Stornaiuolo, A.; Cafiero, M.; Faiella, A.; Simeone, A.
Genome 31, 745-756, 1989
A;Title: Organization of human class I homeobox genes.
A;Reference number: S15036; MUID:90215256; PMID:2576652
A;Accession: S15538
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-66 <BON>
C;Genetics:
A;Gene: GDB:HOXA6
A;Cross-references: GDB:120648; OMIM:142951
A;Map position: 7p15.3-7p15.3
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

  Query Match             96.7%;  Score 89;  DB 2;  Length 66;
  Best Local Similarity   93.8%;  Pred. No. 6.8e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         43 RQIKIWFQNRRMKWKK 58


RESULT 7
JC1161
homeotic protein 3.4 - eastern newt (fragment)
```

C;Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 23-May-1997
C;Accession: JC1161
R;Belleville, S.; Beauchemin, M.; Tremblay, M.; Noiseux, N.; Savard, P.
Gene 114, 179-186, 1992
A;Title: Homeobox-containing genes in the newt are organized in clusters similar to other vertebrates.
A;Reference number: JC1161; MUID:92290273; PMID:1351019
A;Accession: JC1161
A;Molecule type: DNA
A;Residues: 1-71 <BEL>
A;Cross-references: GB:M84001
C;Genetics:
A;Gene: NvHBox-3.4
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;5-61/Domain: homeobox homology <HOX>

  Query Match 96.7%; Score 89; DB 2; Length 71;
  Best Local Similarity 93.8%; Pred. No. 7.3e-07;
  Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
      :||||||||||||||||
Db 46 RQIKIWFQNRRMKWKK 61

RESULT 8
A60084
homeotic protein Hox 3.4 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 07-May-1999
C;Accession: A60084
R;Gaunt, S.J.; Coletta, P.L.; Pravtcheva, D.; Sharpe, P.T.
Development 109, 329-339, 1990
A;Title: Mouse Hox-3.4: homeobox sequence and embryonic expression patterns compared with other members of the Hox gene network.
A;Reference number: A60084; MUID:90382249; PMID:1976088
A;Accession: A60084
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-71 <GAU>
C;Genetics:
A;Map position: 15
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;5-61/Domain: homeobox homology <HOX>

  Query Match 96.7%; Score 89; DB 2; Length 71;
  Best Local Similarity 93.8%; Pred. No. 7.3e-07;
  Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
      :||||||||||||||||
Db 46 RQIKIWFQNRRMKWKK 61

RESULT 9
D34510
homeotic protein H90 - honeybee (fragment)
C;Species: Apis mellifera (honeybee)
C;Date: 22-Jun-1990 #sequence_revision 09-Oct-1992 #text_change 24-Sep-1999
C;Accession: D34510
R;Walldorf, U.; Fleig, R.; Gehring, W.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 9971-9975, 1989
A;Title: Comparison of homeobox-containing genes of the honeybee and Drosophila.
A;Reference number: A34510; MUID:90099384; PMID:2574865
A;Accession: D34510
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-74 <WAL>
A;Cross-references: GB:M29493; NID:g155675; PIDN:AAA27728.1; PID:g155676
A;Note: the authors mistranslated the codons for residues 68-74
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;9-65/Domain: homeobox homology <HOX>

  Query Match 96.7%; Score 89; DB 2; Length 74;
  Best Local Similarity 93.8%; Pred. No. 7.6e-07;
  Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
      :||||||||||||||||
Db 50 RQIKIWFQNRRMKWKK 65

RESULT 10
I51341
homeo box protein - Atlantic salmon (fragment)
C;Species: Salmo salar (Atlantic salmon)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C;Accession: I51341
R;Fjose, A.; Molven, A.; Eiken, H.G.
Gene 62, 141-152, 1988
A;Title: Molecular cloning and characterization of homeobox-containing genes from Atlantic salmon.
A;Reference number: I51341; MUID:88226009; PMID:2897318
A;Accession: I51341
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-75 <FJO>
A;Cross-references: GB:M18903; NID:g213797; PIDN:AAA49559.1; PID:g213798
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

  Query Match 96.7%; Score 89; DB 2; Length 75;
  Best Local Similarity 93.8%; Pred. No. 7.7e-07;
  Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KQIKIWFQNRRMKWKK 16
      :||||||||||||||||
Db 43 RQIKIWFQNRRMKWKK 58

RESULT 11
S58852
homeotic protein Scr homolog - Junonia coenia (fragment)
N;Alternate names: sex combs reduced homeodomain protein
C;Species: Junonia coenia
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 24-Sep-1999
C;Accession: S58852
R;Warren, R.W.; Nagy, L.; Selegue, J.; Gates, J.; Carroll, S.
Nature 372, 458-461, 1994
A;Title: Evolution of homeotic gene regulation and function in flies and
butterflies.
A;Reference number: S58850; MUID:95075456; PMID:7840822
A;Accession: S58852
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 1-75 <WAR>
A;Cross-references: EMBL:L42136; NID:g806495; PIDN:AAA68462.1; PID:g806496
A;Note: the nuclectide sequence was submitted to the EMBL Data Library, May 1995
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;1-40/Domain: homeobox homology (fragment) <HOX>

  Query Match   96.7%; Score 89; DB 2; Length 75;
  Best Local Similarity   93.8%; Pred. No. 7.7e-07;
  Matches   15; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db   25 RQIKIWFQNRRMKWKK 40

RESULT 12
C43559
homeotic protein R3 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Dec-1992 #sequence_revision 01-Dec-1992 #text_change 24-Sep-1999
C;Accession: C43559
R;Falzon, M.; Chung, S.Y.
Development 103, 601-610, 1988
A;Title: The expression of rat homeobox-containing genes is developmentally
regulated and tissue specific.
A;Reference number: A43559; MUID:89231502; PMID:2907739
A;Accession: C43559
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-76 <FAL>
A;Cross-references: GB:M37567; NID:g204634; PIDN:AAA41343.1; PID:g204635
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;20-76/Domain: homeobox homology (fragment) <HOX>

  Query Match   96.7%; Score 89; DB 2; Length 76;
  Best Local Similarity   93.8%; Pred. No. 7.8e-07;
  Matches   15; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db   53 RQIKIWFQNRRMKWKK 68

RESULT 13
I51342
homeo box protein - Atlantic salmon (fragment)
C;Species: Salmo salar (Atlantic salmon)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Sep-1999
C;Accession: I51342
R;Fjose, A.; Molven, A.; Eiken, H.G.
Gene 62, 141-152, 1988
A;Title: Molecular cloning and characterization of homeobox-containing genes
from Atlantic salmon.
A;Reference number: I51341; MUID:88226009; PMID:2897318
A;Accession: I51342
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-78 <FJO>
A;Cross-references: GB:M18904; NID:g213799; PIDN:AAA49560.1; PID:g213800
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

  Query Match   96.7%; Score 89; DB 2; Length 78;
  Best Local Similarity   93.8%; Pred. No. 8e-07;
  Matches   15; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db   43 RQIKIWFQNRRMKWKK 58

RESULT 14
S47605
homeotic protein Hox-7 - Florida lancelet (fragment)
C;Species: Branchiostoma floridae (Florida lancelet)
C;Date: 01-Feb-1995 #sequence_revision 26-May-1995 #text_change 24-Sep-1999
C;Accession: S47605
R;Garcia-Fernandez, J.; Holland, P.W.H.
Nature 370, 563-566, 1994
A;Title: Archetypal organization of the amphioxus Hox gene cluster.
A;Reference number: S47599; MUID:94329179; PMID:7914353
A;Accession: S47605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-81 <GAR>
A;Cross-references: EMBL:Z35147; NID:g520617; PIDN:CAA84519.1; PID:g520618
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;4-60/Domain: homeobox homology <HOX>

  Query Match   96.7%; Score 89; DB 2; Length 81;
  Best Local Similarity   93.8%; Pred. No. 8.3e-07;
  Matches   15; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQIKIWFQNRRMKWKK 16

```
                :|||||||||||||||
Db          45 RQIKIWFQNRRMKWKK 60
```

RESULT 15
B29585
homeotic protein Hox 2.2 precursor - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 15-Dec-1988 #sequence_revision 30-Sep-1991 #text_change 17-Oct-1997
C;Accession: B29585
R;Lonai, P.; Arman, E.; Czosnek, H.; Ruddle, F.H.; Blatt, C.
DNA 6, 409-418, 1987
A;Title: New murine homeoboxes: structure, chromosomal assignment, and differential expression in adult erythropoiesis.
A;Reference number: A29585; MUID:88054465; PMID:2890503
A;Accession: B29585
A;Molecule type: DNA
A;Residues: 1-81 <LON>
A;Cross-references: GB:M18167
A;Note: the authors translated the codon CAG for residue 69 as Glu
C;Superfamily: unassigned homeobox proteins; homeobox homology
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;4-60/Domain: homeobox homology <HOX>

  Query Match   96.7%; Score 89; DB 2; Length 81;
  Best Local Similarity 93.8%; Pred. No. 8.3e-07;
  Matches   15; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db          45 RQIKIWFQNRRMKWKK 60
```

Search completed: October 4, 2004, 18:58:26
Job time : 24.6522 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:44:39 ; Search time 13.2174 Seconds
(without alignments)
63.032 Million cell updates/sec

Title: US-10-048-209-5
Perfect score: 92
Sequence: 1 KQIKIWFQNRRMKWKK 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	89	96.7	48	1	HXB6_XENLA	P31256 xenopus lae
2	89	96.7	49	1	HXA5_SHEEP	Q28599 ovis aries
3	89	96.7	71	1	HXA7_SHEEP	Q28600 ovis aries
4	89	96.7	71	1	HXC5_NOTVI	P31262 notophthalm
5	89	96.7	74	1	HM90_APIME	P15860 apis mellif
6	89	96.7	75	1	HMSA_SALSA	P09636 salmo salar
7	89	96.7	76	1	HXC4_RAT	P18865 rattus norv
8	89	96.7	78	1	HXA5_SALSA	P09637 salmo salar
9	89	96.7	80	1	HXA4_LINSA	P81192 lineus sang
10	89	96.7	81	1	HX5L_BRARE	P09013 brachydanio
11	89	96.7	82	1	HXB5_CHICK	P14838 gallus gall
12	89	96.7	84	1	HXB6_CHICK	P14839 gallus gall
13	89	96.7	86	1	SCR_APIME	P15859 apis mellif
14	89	96.7	87	1	HXC5_XENLA	P09020 xenopus lae
15	89	96.7	93	1	HXB8_PIG	P09078 sus scrofa
16	89	96.7	96	1	HXC6_BRARE	P15862 brachydanio
17	89	96.7	105	1	HXA7_RAT	P09634 rattus norv

18	89	96.7	105	1	HXB4_BRARE	P22574 brachydanio
19	89	96.7	112	1	HXB7_RAT	P18864 rattus norv
20	89	96.7	148	1	HXA5_AMBME	P50208 ambystoma m
21	89	96.7	153	1	HXC6_SHEEP	P49925 ovis aries
22	89	96.7	208	1	HXA7_HETFR	Q9ia25 heterodontu
23	89	96.7	209	1	HXA7_XENLA	P09071 xenopus lae
24	89	96.7	217	1	HXB7_BOVIN	Q9tt89 bos taurus
25	89	96.7	217	1	HXB7_HUMAN	P09629 homo sapien
26	89	96.7	217	1	HXB7_MOUSE	P09024 mus musculu
27	89	96.7	220	1	HB7A_XENLA	Q91771 xenopus lae
28	89	96.7	220	1	HB7B_XENLA	P04476 xenopus lae
29	89	96.7	222	1	HXC5_HUMAN	Q00444 homo sapien
30	89	96.7	222	1	HXC5_MOUSE	P32043 mus musculu
31	89	96.7	224	1	HXB6_HUMAN	P17509 homo sapien
32	89	96.7	224	1	HXB6_MOUSE	P09023 mus musculu
33	89	96.7	225	1	HXA7_MORSA	Q9pwd4 morone saxa
34	89	96.7	228	1	HXB6_BRARE	P15861 brachydanio
35	89	96.7	229	1	HXA6_HETFR	Q9ia24 heterodontu
36	89	96.7	229	1	HXA7_MOUSE	P02830 mus musculu
37	89	96.7	230	1	HXA7_HUMAN	P31268 homo sapien
38	89	96.7	230	1	HXB5_XENLA	P09019 xenopus lae
39	89	96.7	232	1	HXA6_MOUSE	P09092 mus musculu
40	89	96.7	232	1	HXB4_XENLA	P09070 xenopus lae
41	89	96.7	232	1	HXC5_BRARE	P09074 brachydanio
42	89	96.7	233	1	HXA5_RAT	P52949 rattus norv
43	89	96.7	233	1	HXA6_HUMAN	P31267 homo sapien
44	89	96.7	234	1	HXC6_NOTVI	P14858 notophthalm
45	89	96.7	234	1	HXC6_XENLA	P02832 xenopus lae

ALIGNMENTS

```
RESULT 1
HXB6_XENLA
ID   HXB6_XENLA     STANDARD;      PRT;    48 AA.
AC   P31256;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-B6 (XlHox-2.2) (Fragment).
GN   HOXB6 OR XLHOX-2.2.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93043517; PubMed=1384809;
RA   Leroy P., de Robertis E.M.;
RT   "Effects of lithium chloride and retinoic acid on the expression of
RT   genes from the Xenopus laevis Hox 2 complex.";
RL   Dev. Dyn. 194:21-32(1992).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M91587; AAA49750.1; -.
DR   PIR; I51439; I51439.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND     <1     29       HOMEOBOX.
SQ   SEQUENCE   48 AA;  5716 MW;  BC39E36822EDDD2A CRC64;

  Query Match           96.7%;  Score 89;  DB 1;  Length 48;
  Best Local Similarity   93.8%;  Pred. No. 7.5e-08;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         12 RQIKIWFQNRRMKWKK 27


RESULT 2
HXA5_SHEEP
ID   HXA5_SHEEP     STANDARD;      PRT;    49 AA.
AC   Q28599;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-A5 (Fragment).
GN   HOXA5 OR HOXA-5.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Roche P.J.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
```

```
CC       specific positional identities on the anterior-posterior axis.
CC       Also binds to its own promoter. Binds specifically to the motif:
CC       5'-CYYNATTA[TG]Y-3'.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U61978; AAB04754.1; -.
DR   HSSP; P02833; 1HOM.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND     <1     49       HOMEOBOX.
FT   NON_TER      49     49
SQ   SEQUENCE   49 AA;  6331 MW;  1EE702315E7C099B CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 49;
  Best Local Similarity   93.8%;  Pred. No. 7.7e-08;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         32 RQIKIWFQNRRMKWKK 47


RESULT 3
HXA7_SHEEP
ID   HXA7_SHEEP     STANDARD;      PRT;    71 AA.
AC   Q28600;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-A7 (Fragment).
GN   HOXA7 OR HOXA-7.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Roche P.J.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
```

```
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; U61979; AAB04755.1; -.
DR   HSSP; P02833; 9ANT.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND      4     63       HOMEOBOX.
FT   NON_TER      71     71
SQ   SEQUENCE   71 AA;  8888 MW;  931049FAC1BAACB7 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 71;
  Best Local Similarity   93.8%;  Pred. No. 1.1e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         46 RQIKIWFQNRRMKWKK 61


RESULT 4
HXC5_NOTVI
ID   HXC5_NOTVI     STANDARD;      PRT;    71 AA.
AC   P31262;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-C5 (NvHbox-3.4) (Fragment).
OS   Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;
OC   Notophthalmus.
OX   NCBI_TaxID=8316;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92290273; PubMed=1351019;
```

```
RA   Belleville S., Beauchemin M., Tremblay M., Noiseux N., Savard P.;
RT   "Homeobox-containing genes in the newt are organized in clusters
RT   similar to other vertebrates.";
RL   Gene 114:179-186(1992).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M84001; AAA49397.1; ALT_INIT.
DR   PIR; JC1161; JC1161.
DR   HSSP; P02833; 1SAN.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND      4     63       HOMEOBOX.
FT   NON_TER      71     71
SQ   SEQUENCE   71 AA;  8979 MW;  07999FDE89995B42 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 71;
  Best Local Similarity   93.8%;  Pred. No. 1.1e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         46 RQIKIWFQNRRMKWKK 61


RESULT 5
HM90_APIME
ID   HM90_APIME     STANDARD;      PRT;    74 AA.
AC   P15860;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Homeobox protein H90 (Fragment).
OS   Apis mellifera (Honeybee).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC   Apidae; Apis.
OX   NCBI_TaxID=7460;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90099384; PubMed=2574865;
RA   Walldorf U., Fleig R., Gehring W.J.;
RT   "Comparison of homeobox-containing genes of the honeybee and
RT   Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:9971-9975(1989).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   -!- SIMILARITY: Contains 1 homeobox domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M29493; AAA27728.1; -.
DR   PIR; D34510; D34510.
DR   HSSP; P02833; 1HOM.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT   NON_TER       1      1
FT   DNA_BIND      8     67       HOMEOBOX.
FT   NON_TER      74     74
SQ   SEQUENCE   74 AA;  9263 MW;  5FC8FB4F723D3837 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 74;
  Best Local Similarity   93.8%;  Pred. No. 1.2e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         50 RQIKIWFQNRRMKWKK 65


RESULT 6
HMSA_SALSA
ID   HMSA_SALSA     STANDARD;      PRT;    75 AA.
AC   P09636;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Homeobox protein S12-A (Fragment).
OS   Salmo salar (Atlantic salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
```

```
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX   NCBI_TaxID=8030;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88226009; PubMed=2897318;
RA   Fjose A., Molven A., Eiken H.G.;
RT   "Molecular cloning and characterization of homeo-box-containing genes
RT   from Atlantic salmon.";
RL   Gene 62:141-152(1988).
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   -!- SIMILARITY: Contains 1 homeobox domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M18903; AAA49559.1; -.
DR   PIR; I51341; I51341.
DR   HSSP; P02833; 9ANT.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT   NON_TER       1      1
FT   DNA_BIND      1     60       HOMEOBOX.
FT   NON_TER      75     75
SQ   SEQUENCE   75 AA;  9330 MW;  FC02C3672F35475D CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 75;
  Best Local Similarity   93.8%;  Pred. No. 1.2e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         43 RQIKIWFQNRRMKWKK 58


RESULT 7
HXC4_RAT
ID   HXC4_RAT       STANDARD;      PRT;    76 AA.
AC   P18865;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Homeobox protein Hox-C4 (R3) (Fragment).
GN   HOXC4 OR HOXC-4.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=89231502; PubMed=2907739;
RA   Falzon M., Chung S.Y.;
RT   "The expression of rat homeobox-containing genes is developmentally
RT   regulated and tissue specific.";
RL   Development 103:601-610(1988).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Predominantly spinal cord and kidney.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC       subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M37567; AAA41343.1; -.
DR   PIR; C43559; C43559.
DR   HSSP; P02833; 9ANT.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND     11     70       HOMEOBOX.
SQ   SEQUENCE   76 AA;  9293 MW;  5235F665C0672385 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 76;
  Best Local Similarity   93.8%;  Pred. No. 1.2e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         53 RQIKIWFQNRRMKWKK 68


RESULT 8
HXA5_SALSA
```

```
ID   HXA5_SALSA     STANDARD;      PRT;    78 AA.
AC   P09637;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-A5 (S12-B) (Fragment).
GN   HOXA5.
OS   Salmo salar (Atlantic salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX   NCBI_TaxID=8030;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88226009; PubMed=2897318;
RA   Fjose A., Molven A., Eiken H.G.;
RT   "Molecular cloning and characterization of homeo-box-containing genes
RT   from Atlantic salmon.";
RL   Gene 62:141-152(1988).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M18904; AAA49560.1; -.
DR   PIR; I51342; I51342.
DR   HSSP; P02833; 9ANT.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND      1     60       HOMEOBOX.
SQ   SEQUENCE   78 AA;  9489 MW;  828DEBDDF78AC820 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 78;
  Best Local Similarity   93.8%;  Pred. No. 1.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
```

```
Db         43 RQIKIWFQNRRMKWKK 58


RESULT 9
HXA4_LINSA
ID   HXA4_LINSA     STANDARD;      PRT;    80 AA.
AC   P81192;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Homeobox protein Hox-A4 (LsHox 4) (Fragment).
GN   HOXA4.
OS   Lineus sanguineus (Ribbon worm).
OC   Eukaryota; Metazoa; Nemertea; Anopla; Heteronemertea; Lineidae;
OC   Lineus.
OX   NCBI_TaxID=48190;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98169491; PubMed=9501210;
RA   Kmita-Cunisse M., Locsli F., Bierne J., Gehring W.J.;
RT   "Homeobox genes in the ribbonworm Lineus sanguineus: evolutionary
RT   implications.";
RL   Proc. Natl. Acad. Sci. U.S.A. 95:3030-3035(1998).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: Belongs to the Antp homeobox family. Deformed
CC       subfamily.
DR   HSSP; P02833; 9ANT.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND     21     80       HOMEOBOX.
FT   NON_TER      80     80
SQ   SEQUENCE   80 AA;  9860 MW;  F2CE1B01CB8042F1 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 80;
  Best Local Similarity   93.8%;  Pred. No. 1.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         53 RQIKIWFQNRRMKWKK 68
```

```
RESULT 10
HX5L_BRARE
ID   HX5L_BRARE     STANDARD;      PRT;    81 AA.
AC   P09013;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-B5 like (ZF-54) (Fragment).
GN   HOXB5B OR ZF54 OR ZF-54.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89016617; PubMed=2902580;
RA   Njolstad P.R., Molven A., Hordvik I., Apold J., Fjose A.;
RT   "Primary structure, developmentally regulated expression and
RT   potential duplication of the zebrafish homeobox gene ZF-21.";
RL   Nucleic Acids Res. 16:9097-9113(1988).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X12803; CAA31291.1; -.
DR   HSSP; P02833; 1SAN.
DR   ZFIN; ZDB-GENE-000823-6; hoxb5b.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND      6     65       HOMEOBOX.
SQ   SEQUENCE   81 AA;  9977 MW;  B7698AEFFEB3C6B4 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 81;
  Best Local Similarity   93.8%;  Pred. No. 1.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         48 RQIKIWFQNRRMKWKK 63


RESULT 11
HXB5_CHICK
ID   HXB5_CHICK     STANDARD;      PRT;    82 AA.
AC   P14838;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-B5 (Ghox-2.1) (Fragment).
GN   HOXB5 OR GHOX-2.1.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythrocyte;
RX   MEDLINE=90126373; PubMed=2575515;
RA   Wedden S.E., Pang K., Eichele G.;
RT   "Expression pattern of homeobox-containing genes during chick
RT   embryogenesis.";
RL   Development 105:639-650(1989).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X16846; CAA34743.1; -.
DR   PIR; S08302; S08302.
DR   HSSP; P02833; 1SAN.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND      7     66       HOMEOBOX.
SQ   SEQUENCE   82 AA;  9877 MW;  53F70ACDC9FDEF8F CRC64;
```

```
  Query Match             96.7%;  Score 89;  DB 1;  Length 82;
  Best Local Similarity   93.8%;  Pred. No. 1.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         49 RQIKIWFQNRRMKWKK 64


RESULT 12
HXB6_CHICK
ID   HXB6_CHICK     STANDARD;      PRT;    84 AA.
AC   P14839;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-B6 (GHOX-2.2) (Fragment).
GN   HOXB6 OR GHOX-2.2.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Erythrocyte;
RX   MEDLINE=90126373; PubMed=2575515;
RA   Wedden S.E., Pang K., Eichele G.;
RT   "Expression pattern of homeobox-containing genes during chick
RT   embryogenesis.";
RL   Development 105:639-650(1989).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X16847; CAA34744.1; -.
DR   PIR; S08303; S08303.
DR   HSSP; P02833; 1HOM.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND      7     66       HOMEOBOX.
SQ   SEQUENCE   84 AA;  10279 MW;  BC06B10165B19E71 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 84;
  Best Local Similarity   93.8%;  Pred. No. 1.4e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         49 RQIKIWFQNRRMKWKK 64


RESULT 13
SCR_APIME
ID   SCR_APIME      STANDARD;      PRT;    86 AA.
AC   P15859;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein H55 (Fragment).
OS   Apis mellifera (Honeybee).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC   Apidae; Apis.
OX   NCBI_TaxID=7460;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90099384; PubMed=2574865;
RA   Walldorf U., Fleig R., Gehring W.J.;
RT   "Comparison of homeobox-containing genes of the honeybee and
RT   Drosophila.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:9971-9975(1989).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE ANTP HOMEOBOX FAMILY. STRONGEST, TO SCR
CC       OF DROSOPHILA.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M29488; AAA27723.1; -.
DR   PIR; A34510; A34510.
DR   HSSP; P02833; 1SAN.
DR   InterPro; IPR001827; Antennapedia.
```

```
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT   NON_TER       1      1
FT   DNA_BIND      8     67       HOMEOBOX.
FT   NON_TER      86     86
SQ   SEQUENCE   86 AA;  10713 MW;  2A49AB857C138AB8 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 86;
  Best Local Similarity   93.8%;  Pred. No. 1.4e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         50 RQIKIWFQNRRMKWKK 65


RESULT 14
HXC5_XENLA
ID   HXC5_XENLA     STANDARD;      PRT;    87 AA.
AC   P09020;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-C5 (XlHbox-5) (Fragment).
GN   HOXC5 OR XLHBOX5.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88157707; PubMed=2894634;
RA   Fritz A., de Robertis E.M.;
RT   "Xenopus homeobox-containing cDNAs expressed in early development.";
RL   Nucleic Acids Res. 16:1453-1469(1988).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED EXCLUSIVELY IN EARLY EMBRYOS.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X07105; CAA30126.1; -.
DR   PIR; S00589; S00589.
DR   HSSP; P02833; 9ANT.
DR   TRANSFAC; T03765; -.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00025; ANTENNAPEDIA.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   SITE          5     10       ANTP-TYPE HEXAPEPTIDE.
FT   DNA_BIND     20     79       HOMEOBOX.
SQ   SEQUENCE   87 AA;  11058 MW;  E67939E334E2BA43 CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 87;
  Best Local Similarity   93.8%;  Pred. No. 1.4e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         62 RQIKIWFQNRRMKWKK 77


RESULT 15
HXB8_PIG
ID   HXB8_PIG       STANDARD;      PRT;    93 AA.
AC   P09078;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homeobox protein Hox-B8 (Hox-2.4) (Fragment).
GN   HOXB8 OR HOX-2.4.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89057478; PubMed=2904133;
RA   Miller J.R., Gaunt S.J., Sharpe P.T.;
RT   "Pig Hox-2.4 has accumulated a frameshift mutation relative to mouse
RT   Hox-2.4.";
RL   Nucleic Acids Res. 16:10364-10364(1988).
CC   -!- FUNCTION: Sequence-specific transcription factor which is part of
CC       a developmental regulatory system that provides cells with
CC       specific positional identities on the anterior-posterior axis.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: Belongs to the Antp homeobox family.
CC   ----------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X06668; CAB57825.1; ALT_SEQ.
DR   HSSP; P02834; 1B8I.
DR   InterPro; IPR001827; Antennapedia.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00032; ANTENNAPEDIA; PARTIAL.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW   Transcription regulation.
FT   NON_TER       1      1
FT   DNA_BIND      7     66       HOMEOBOX.
SQ   SEQUENCE   93 AA;  10872 MW;  7841DD6D17634EEC CRC64;

  Query Match             96.7%;  Score 89;  DB 1;  Length 93;
  Best Local Similarity   93.8%;  Pred. No. 1.5e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         49 RQIKIWFQNRRMKWKK 64


Search completed: October  4, 2004, 18:57:19
Job time : 14.2174 secs
```

```
                                GenCore version 5.1.6
                          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          October  4, 2004, 18:37:37 ; Search time 75.8261 Seconds
                                             (without alignments)
                                             66.577 Million cell updates/sec

Title:           US-10-048-209-5
Perfect score:   92
Sequence:        1 KQIKIWFQNRRMKWKK 16

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:      1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        SPTREMBL_25:*
                 1:  sp_archea:*
                 2:  sp_bacteria:*
                 3:  sp_fungi:*
                 4:  sp_human:*
                 5:  sp_invertebrate:*
                 6:  sp_mammal:*
                 7:  sp_mhc:*
                 8:  sp_organelle:*
                 9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_virus:*
                 13:  sp_vertebrate:*
                 14:  sp_unclassified:*
                 15:  sp_rvirus:*
                 16:  sp_bacteriap:*
                 17:  sp_archeap:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                        SUMMARIES

                                  %
Result                            Query
  No.    Score                    Match Length DB  ID                          Description
-------------------------------------------------------------------------------------------
```

1	90	97.8	190	5	Q9NLA0	Q9nla0 podocoryne
2	89	96.7	33	5	Q86FU0	Q86fu0 drosophila
3	89	96.7	39	13	O57368	O57368 brachydanio
4	89	96.7	42	11	Q80WH6	Q80wh6 rattus sp.
5	89	96.7	43	13	O57359	O57359 brachydanio
6	89	96.7	46	13	Q9PVR9	Q9pvr9 oryzias lat
7	89	96.7	51	5	Q27413	Q27413 ctenodrilus
8	89	96.7	51	5	Q23743	Q23743 ctenodrilus
9	89	96.7	51	5	Q26407	Q26407 ctenodrilus
10	89	96.7	57	13	Q9PVR8	Q9pvr8 oryzias lat
11	89	96.7	58	5	Q9Y188	Q9y188 priapulus c
12	89	96.7	58	5	Q25208	Q25208 junonia coe
13	89	96.7	58	13	O57362	O57362 brachydanio
14	89	96.7	59	5	Q8WRM9	Q8wrm9 lithobius a
15	89	96.7	59	5	Q9NB42	Q9nb42 anopheles g
16	89	96.7	59	13	Q9PVR5	Q9pvr5 oryzias lat
17	89	96.7	60	5	O77143	O77143 archegozete
18	89	96.7	60	5	O77139	O77139 archegozete
19	89	96.7	60	11	Q80WH7	Q80wh7 rattus sp.
20	89	96.7	60	11	Q80WH4	Q80wh4 rattus sp.
21	89	96.7	60	13	Q8QGL5	Q8qgl5 petromyzon
22	89	96.7	60	13	Q8QGL3	Q8qgl3 petromyzon
23	89	96.7	60	13	Q8QGL6	Q8qgl6 petromyzon
24	89	96.7	60	13	Q8QGL2	Q8qgl2 petromyzon
25	89	96.7	60	13	Q8QGL8	Q8qgl8 petromyzon
26	89	96.7	60	13	Q8QGL7	Q8qgl7 petromyzon
27	89	96.7	61	5	Q27910	Q27910 polyandroca
28	89	96.7	63	5	O77138	O77138 archegozete
29	89	96.7	63	5	Q8MXB2	Q8mxb2 holopneuste
30	89	96.7	66	13	O57356	O57356 brachydanio
31	89	96.7	69	5	Q9U9T4	Q9u9t4 nereis vire
32	89	96.7	69	5	Q9BMF7	Q9bmf7 haliotis as
33	89	96.7	70	5	Q967W5	Q967w5 folsomia ca
34	89	96.7	70	13	Q801B4	Q801b4 latimeria m
35	89	96.7	71	13	Q9PVS3	Q9pvs3 oryzias lat
36	89	96.7	71	13	Q9PVS1	Q9pvs1 oryzias lat
37	89	96.7	73	5	Q9Y186	Q9y186 priapulus c
38	89	96.7	73	5	Q86D93	Q86d93 spadella ce
39	89	96.7	74	13	O57367	O57367 brachydanio
40	89	96.7	75	5	Q25209	Q25209 junonia coe
41	89	96.7	75	13	Q9PVR6	Q9pvr6 oryzias lat
42	89	96.7	76	5	O44257	O44257 ethmostigmu
43	89	96.7	76	5	Q86NB1	Q86nb1 calanus hel
44	89	96.7	77	5	O44260	O44260 ethmostigmu
45	89	96.7	77	5	Q9Y187	Q9y187 priapulus c

ALIGNMENTS

RESULT 1
Q9NLA0
ID   Q9NLA0   PRELIMINARY;   PRT;   190 AA.
AC   Q9NLA0;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cnox2-Pc.
GN   CNOX2-PC.
OS   Podocoryne carnea.
OC   Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC   Hydractiniidae; Podocoryne.
OX   NCBI_TaxID=6096;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21110504; PubMed=11180816;
RA   Masuda-Nakagawa L.M., Groeger H., Aerne B.L., Schmid V.;
RT   "The HOX-like gene Cnox2-Pc is expressed at the anterior region in all
RT   life cycle stages of the jellyfish Podocoryne carnea.";
RL   Dev. Genes Evol. 210:151-156(2000).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AB014684; BAA94091.1; -.
DR   HSSP; P14653; 1B72.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
SQ   SEQUENCE   190 AA; 22929 MW; A476A14D8147288A CRC64;

  Query Match   97.8%; Score 90; DB 5; Length 190;
  Best Local Similarity   87.5%; Pred. No. 1.6e-06;
  Matches   14; Conservative   2; Mismatches   0; Indels   0; Gaps   0;

Qy   1 KQIKIWFQNRRMKWKK 16
              ||:|:|||||||||||
Db   161 KQVKVWFQNRRMKWKK 176

RESULT 2
Q86FU0
ID   Q86FU0   PRELIMINARY;   PRT;   33 AA.
AC   Q86FU0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Antennapedia complex (Fragment).
GN   ANT-C.
OS   Drosophila pseudoobscura (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7237;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93292933; PubMed=8099892;

```
RA   Randazzo F.M., Seeger M.A., Huss C.A., Sweeney M.A., Cecil J.K.,
RA   Kaufman T.C.;
RT   "Structural changes in the antennapedia complex of Drosophila
RT   pseudoobscura.";
RL   Genetics 134:319-330(1993).
DR   EMBL; S63455; AAP13946.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
FT   NON_TER       1      1
FT   NON_TER      33     33
SQ   SEQUENCE   33 AA;  3963 MW;  D78E37ED81FD45DF CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 33;
  Best Local Similarity   93.8%;  Pred. No. 4.1e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db          7 RQIKIWFQNRRMKWKK 22


RESULT 3
O57368
ID   O57368      PRELIMINARY;      PRT;    39 AA.
AC   O57368;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hoxc5 protein (Fragment).
GN   HOXC5A OR HOXC5.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Prince V.E., Joly L., Ekker M., Ho R.K.;
RT   "Zebrafish hox genes: genomic organization and modified colinear
RT   expression patterns in the trunk.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; Y14539; CAA74874.1; -.
DR   ZFIN; ZDB-GENE-980526-533; hoxc5a.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   39 AA;  4827 MW;  592A0FEC12E58860 CRC64;

  Query Match             96.7%;  Score 89;  DB 13;  Length 39;
  Best Local Similarity   93.8%;  Pred. No. 4.8e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :|||||||||||||||
Db         14 RQIKIWFQNRRMKWKK 29


RESULT 4
Q80WH6
ID   Q80WH6      PRELIMINARY;      PRT;    42 AA.
AC   Q80WH6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hox-A|Hox-1 (Fragment).
GN   HOX-A|HOX-1.
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95217128; PubMed=7702549;
RA   Sakoyama Y., Mizuta I., Ogasawara N., Yoshikawa H.;
RT   "Cloning of rat homeobox genes.";
RL   Biochem. Genet. 32:351-360(1994).
DR   EMBL; S76290; AAP31864.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
FT   NON_TER       1      1
FT   NON_TER      42     42
SQ   SEQUENCE   42 AA;  5494 MW;  38E5153B92216FE9 CRC64;

  Query Match             96.7%;  Score 89;  DB 11;  Length 42;
  Best Local Similarity   93.8%;  Pred. No. 5.2e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         25 RQIKIWFQNRRMKWKK 40


RESULT 5
O57359
ID   O57359      PRELIMINARY;      PRT;    43 AA.
AC   O57359;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hoxa5 protein (Fragment).
GN   HOXB5B OR HOXA5.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Prince V.E., Joly L., Ekker M., Ho R.K.;
RT   "Zebrafish hox genes: genomic organization and modified colinear
RT   expression patterns in the trunk.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; Y14526; CAA74861.1; -.
DR   ZFIN; ZDB-GENE-000823-6; hoxb5b.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1       1
SQ   SEQUENCE   43 AA;  5050 MW;  53034C37F3DFA596 CRC64;

  Query Match             96.7%;  Score 89;  DB 13;  Length 43;
  Best Local Similarity   93.8%;  Pred. No. 5.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         10 RQIKIWFQNRRMKWKK 25


RESULT 6
Q9PVR9
ID   Q9PVR9      PRELIMINARY;      PRT;    46 AA.
AC   Q9PVR9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   HOXC5A (Fragment).
GN   HOXC5A.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kondo S., Naruse K., Shima A.;
RT   "Hox genes of the medakafish Oryzias latipes.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AB026960; BAA86243.1; -.
DR   HSSP; P02833; 1HOM.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1       1
SQ   SEQUENCE   46 AA;  5955 MW;  6039999ED4294DD3 CRC64;

  Query Match             96.7%;  Score 89;  DB 13;  Length 46;
  Best Local Similarity   93.8%;  Pred. No. 5.7e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         23 RQIKIWFQNRRMKWKK 38


RESULT 7
Q27413
ID   Q27413      PRELIMINARY;      PRT;    51 AA.
AC   Q27413;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   LOX5 ORTHOLOG homeobox (Fragment).
GN   CTS-LOX5.
OS   Ctenodrilus serratus.
OC   Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC   Terebellida; Ctenodrilidae; Ctenodrilus.
OX   NCBI_TaxID=40316;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94356262; PubMed=7915607;
RA   Dick M.H., Buss L.W.;
RT   "A PCR-based survey of homeobox genes in Ctenodrilus serratus
```

```
RT   (Annelida: Polychaeta).";
RL   Mol. Phylogenet. Evol. 3:146-158(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Dick M.H., Buss L.W.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; U26629; AAC46851.1; -.
DR   EMBL; S76226; AAB31777.1; -.
DR   HSSP; P02833; 9ANT.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
FT   NON_TER      51     51

SQ   SEQUENCE   51 AA;  6278 MW;  88C8F65161E94A22 CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 51;
  Best Local Similarity   93.8%;  Pred. No. 6.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         23 RQIKIWFQNRRMKWKK 38


RESULT 8
Q23743
ID   Q23743      PRELIMINARY;      PRT;    51 AA.
AC   Q23743;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Deformed ortholog homeobox (Fragment).
GN   CTS-DFD.
OS   Ctenodrilus serratus.
OC   Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC   Terebellida; Ctenodrilidae; Ctenodrilus.
OX   NCBI_TaxID=40316;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94356262; PubMed=7915607;
RA   Dick M.H., Buss L.W.;
RT   "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RT   (Annelida: Polychaeta).";
RL   Mol. Phylogenet. Evol. 3:146-158(1994).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Dick M.H., Buss L.W.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; U26627; AAC46849.1; -.
DR   HSSP; P02833; 9ANT.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
FT   NON_TER      51     51
SQ   SEQUENCE   51 AA;  6533 MW;  9EDB50C927FBCBD5 CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 51;
  Best Local Similarity   93.8%;  Pred. No. 6.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         23 RQIKIWFQNRRMKWKK 38


RESULT 9
Q26407
ID   Q26407      PRELIMINARY;      PRT;    51 AA.
AC   Q26407;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CTs-Dfd protein (Fragment).
GN   CTS-DFD.
OS   Ctenodrilus serratus.
OC   Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
OC   Terebellida; Ctenodrilidae; Ctenodrilus.
OX   NCBI_TaxID=40316;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94356262; PubMed=7915607;
RA   Dick M.H., Buss L.W.;
RT   "A PCR-based survey of homeobox genes in Ctenodrilus serratus
RT   (Annelida: Polychaeta).";
RL   Mol. Phylogenet. Evol. 3:146-158(1994).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; S76416; AAB31775.1; -.
DR   HSSP; P02833; 9ANT.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
```

```
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   51 AA;  6533 MW;  9EDB50C927FBCBD5 CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 51;
  Best Local Similarity   93.8%;  Pred. No. 6.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         23 RQIKIWFQNRRMKWKK 38


RESULT 10
Q9PVR8
ID   Q9PVR8      PRELIMINARY;      PRT;    57 AA.
AC   Q9PVR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   HOXA5A (Fragment).
GN   HOXA5A.
OS   Oryzias latipes (Medaka fish) (Japanese ricefish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
OX   NCBI_TaxID=8090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kondo S., Naruse K., Shima A.;
RT   "Hox genes of the medakafish Oryzias latipes.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AB026961; BAA86244.1; -.
DR   HSSP; P02833; 1HOM.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   57 AA;  6891 MW;  54A6430320F68C04 CRC64;

  Query Match             96.7%;  Score 89;  DB 13;  Length 57;
  Best Local Similarity   93.8%;  Pred. No. 7e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         23 RQIKIWFQNRRMKWKK 38


RESULT 11
Q9Y188
ID   Q9Y188      PRELIMINARY;      PRT;    58 AA.
AC   Q9Y188;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   HB1 homeodomain protein (Fragment).
GN   HB1.
OS   Priapulus caudatus.
OC   Eukaryota; Metazoa; Priapulida; Priapulidae; Priapulus.
OX   NCBI_TaxID=37621;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99318125; PubMed=10391241;
RA   de Rosa R., Grenier J.K., Andreeva T., Cook C.E., Adoutte A., Akam M.,
RA   Carroll S.B., Balavoine G.;
RT   "Hox genes in brachiopods and priapulids and protostome evolution.";
RL   Nature 399:772-776(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AF144868; AAD40644.1; -.
DR   HSSP; P02833; 9ANT.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   58 AA;  7323 MW;  572F30DA57C9A613 CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 58;
  Best Local Similarity   93.8%;  Pred. No. 7.1e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         24 RQIKIWFQNRRMKWKK 39
```

```
RESULT 12
Q25208
ID   Q25208      PRELIMINARY;      PRT;    58 AA.
AC   Q25208;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Antennapedia protein (Fragment).
GN   ANTENNAPEDIA.
OS   Junonia coenia (Peacock butterfly) (Precis coenia).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OX   NCBI_TaxID=39708;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95075456; PubMed=7840822;
RA   Warren R.W., Nagy L., Selegue J., Gates J., Carroll S.;
RT   "Evolution of homeotic gene regulation and function in flies and
RT   butterflies.";
RL   Nature 372:458-461(1994).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; L42135; AAA68461.1; -.
DR   PIR; S58850; S58850.
DR   HSSP; P02833; 1HOM.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
FT   NON_TER      58     58
SQ   SEQUENCE   58 AA;  7583 MW;  BD69B4875BAE565E CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 58;
  Best Local Similarity   93.8%;  Pred. No. 7.1e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         43 RQIKIWFQNRRMKWKK 58


RESULT 13
O57362
ID   O57362      PRELIMINARY;      PRT;    58 AA.
AC   O57362;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hoxb7 protein (Fragment).
GN   HOXB7A OR HOXB7.
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Prince V.E., Joly L., Ekker M., Ho R.K.;
RT   "Zebrafish hox genes: genomic organization and modified colinear
RT   expression patterns in the trunk.";
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; Y14533; CAA74868.1; -.
DR   TRANSFAC; T03635; -.
DR   ZFIN; ZDB-GENE-000329-2; hoxb7a.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1      1
SQ   SEQUENCE   58 AA;  6863 MW;  41EA6F14488DA1E5 CRC64;

  Query Match             96.7%;  Score 89;  DB 13;  Length 58;
  Best Local Similarity   93.8%;  Pred. No. 7.1e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         14 RQIKIWFQNRRMKWKK 29


RESULT 14
Q8WRM9
ID   Q8WRM9      PRELIMINARY;      PRT;    59 AA.
AC   Q8WRM9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Antennapedia (Fragment).
GN   ANTP.
OS   Lithobius atkinsoni.
OC   Eukaryota; Metazoa; Arthropoda; Myriapoda; Chilopoda;
OC   Pleurostigmophora; Lithobiomorpha; Lithobiidae; Lithobius.
```

```
OX   NCBI_TaxID=177213;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hughes C.L., Kaufman T.C.;
RT   "Exploring the myriapod body plan: expression patterns of the ten Hox
RT   genes in a centipede.";
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AF434996; AAL36901.1; -.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   InterPro; IPR000047; HTH_lambrepressr.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   PRINTS; PR00031; HTHREPRESSR.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1       1
SQ   SEQUENCE   59 AA;  7093 MW;  9E60036CE0D515C1 CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 59;
  Best Local Similarity   93.8%;  Pred. No. 7.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         23 RQIKIWFQNRRMKWKK 38


RESULT 15
Q9NB42
ID   Q9NB42      PRELIMINARY;      PRT;    59 AA.
AC   Q9NB42;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Transcription factor deformed (Fragment).
GN   DFD.
OS   Anopheles gambiae (African malaria mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX   NCBI_TaxID=7165;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21150895; PubMed=11256376;
RA   Powers T.P., Hogan J., Ke Z., Dymbrowski K., Wang X., Collins F.H.,
RA   Kaufman T.C.;
RT   "Characterization of the Hox cluster from the mosquito Anopheles
RT   gambiae (Diptera: Culicidae).";
RL   Evol. Dev. 2:311-325(2000).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR   EMBL; AF269155; AAF91400.1; -.
DR   HSSP; P02833; 1SAN.
DR   GO; GO:0005634; C:nucleus; IEA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR001356; Homeobox.
DR   Pfam; PF00046; homeobox; 1.
DR   PRINTS; PR00024; HOMEOBOX.
DR   ProDom; PD000010; Homeobox; 1.
DR   SMART; SM00389; HOX; 1.
DR   PROSITE; PS00027; HOMEOBOX_1; 1.
DR   PROSITE; PS50071; HOMEOBOX_2; 1.
KW   DNA-binding; Homeobox; Nuclear protein.
FT   NON_TER       1       1
FT   NON_TER      59      59
SQ   SEQUENCE   59 AA;  7621 MW;  C38A2505A81D9952 CRC64;

  Query Match             96.7%;  Score 89;  DB 5;  Length 59;
  Best Local Similarity   93.8%;  Pred. No. 7.3e-07;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 KQIKIWFQNRRMKWKK 16
              :||||||||||||||||
Db         43 RQIKIWFQNRRMKWKK 58


Search completed: October  4, 2004, 18:56:27
Job time : 78.8261 secs
```

A FIBRIL-BLOCKING PEPTIDE, A METHOD FOR PREVENTING FIBRIL FORMATION

Fred J. Stevens
Yair Argon
David Davis
Rosemarie Raffen